**Signal to Noise: Understanding it, Measuring it, and Improving it (Part 1)**

May 01 2009

You've probably heard it before and if you continue to read my columns here, you'll hear it a hundred more times -- astrophotography is all about signal-to-noise ratios (SNR). But, what does that mean and can such a blanket statement be true? I mean, really. It's all about SNR? What about aperture? What about f-ratio? What about Camera X being better than Camera Y? What about monochrome vs. colour cameras? What about cooled vs. Un-cooled and dedicated astro-cams vs. DSLRs? What about refractors vs. reflectors? What about dark skies vs. urban skies? Well, my answer to all of those is that yes, they all matter but that they can all be distilled down to some form of SNR (albeit sometimes a bit more complex than just a single SNR number).

So, to kick off this column, we're going to talk about SNR, what affects it, and how to measure it. Believe it or not, you can do a lot of tests on your camera and come up with clear, precise, repeatable results. For equipment, you need your camera, a lenscap, and a stack of paper. Optional accessories include a piece of tin foil, a rubber band, and an SLR lens or your telescope (small refractors work nicely here, although aren't as handy as an SLR lens). You don't need fancy software either. While I use Matlab for my tests, you can use the freeware program ImageJ and arrive at the exact same results. Honestly, even if you're a Luddite, you can do it. Before you run the tests, though, you should have some notion of what SNR is and what the various sources of noise are in our images. Knowing that, and knowing how your camera behaves will let you figure out how to get the best images and make the most out of your time under the stars.

In this instalment, we'll cover the basic question: What is SNR? Next time, we'll cover how to measure some of the various kinds of noise in your camera. After that, we'll cover some implications for all this on how you should go about taking your images.

*SNR in a Perfect World*

When we take an image of something, be it M51 or a picture of a child, photons pass through the objective (lens or mirror) and hit the sensor array (CCD or CMOS). The array has a whole bunch of individual sensors (that make the pixels) and each one does the same thing. Each one tries to determine how many photons struck it. This starts off as an analog signal (e.g., a voltage) and gets converted into a number by something called an analog-to-digital converter (ADC). Suppose we're taking an image of a dark gray bar and a light gray bar. We might capture 1,000 photons from the dark gray bar and 2,000 photons from the light gray bar during the time when the shutter is open and photons are allowed to pass through the objective and to the sensor. In a perfect world, each time we did this, we'd get the same 1,000 and 2,000 photons and each time this might read out intensity values of 1,000 and 2,000 in our image. Each time we did this and each pixel that's aimed at the bars would end up with these values. In this perfect world, we have signals of 1,000 and 2,000 and we have no noise at all. There is no variability here from pixel to pixel or from image to image which means that we have no noise.

SNR stands for Signal to Noise Ratio and is simply the ratio of the signal (or, to be picky, often the signal plus noise) to the noise. So, we have 1,000 / 0 for the SNR on the dark gray bars and 2,000 / 0 for the SNR on the light gray bars. In both cases, the SNR is infinite. We have a perfect image in this perfect world.

*Noise is a Reality*

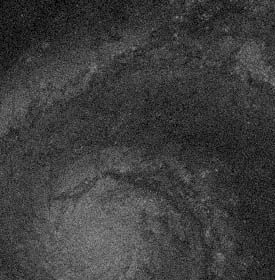
Reality is far from perfect, however, and you will always have noise. In fact, you'll always have several types of noise. We've got read noise, dark current noise, and shot noise (both target and skyglow variants) to name a few. Each of these is going to conspire to keep you from bringing out that faint bit of wispy nebulosity from the background. Before we cover these, though, it's worth going over a few visual examples of SNR. To make these examples, I used the wonderful, ultra-deep shot of M51 from the Hubble Space Telescope, removing the colour, resizing, and cropping to create a very nice, high-SNR image.

|  |  |
| --- | --- |
| https://www.cloudynights.com/images/Fishing-For-Photons/SNR%20Part%201%20SaveAs_html_6d1d0be.jpg | Original image |
| https://www.cloudynights.com/images/Fishing-For-Photons/SNR%20Part%201%20SaveAs_html_m646119eb.jpg | Image with 5% Gaussian noise added |
| https://www.cloudynights.com/images/Fishing-For-Photons/SNR%20Part%201%20SaveAs_html_63a1ef2c.jpg | Image with 10% Gaussain noise added |

On the top, we have the original image and below we have two images in which I've added Gaussian noise. The first one has 5% noise added (using PhotoShop) and the second has 10%. It doesn't take much to see that we'd rather have the top image than the middle and rather have the middle than the bottom. (Although, in truth, if any of my individual frames of M51 looked like the bottom one, I'd be a happy camper!).

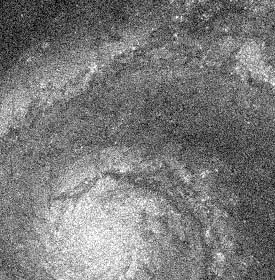
We can even use PhotoShop (or ImageJ) to figure the SNR in parts of this image. I grabbed a small portion of even gray there about centered vertically and about 3/4ths of the way to the right and used PhotoShop to measure the mean signal and standard deviation. The mean is just the average and the standard deviation is a measure of the noise (how much the sampled pixels tend to vary about that mean). In the top image, the mean was 85 and the standard deviation was 6. That makes for an SNR in that area of 14.17 (85 / 6). In the middle image, it was 84.8 and 13.9 for an SNR of 6.10 and in the bottom image it was 84.5 and 26.6 for an SNR of 3.18.

Note that in each of these, the signal is the same. That is, the intensity of the image hasn't changed (it's about 85). What has changed is the amount of noise (the amount of variability around that mean). What happens if we cut the signal down? Well, if we cut the signal in half and add 5% noise like the middle example, we'll end up with something that looks dim and noisy. It'll look worse than the 5% image for sure and some might think it's worse than the 10% image. After all, it's pretty dim. Here's that image.



The mean here is now 42.5 (half of the 85 it was before) and the standard deviation is now 13. The SNR then is 3.27 - almost exactly the same SNR we had in the 10% noise case above. Why? This looks like a crummy image and it looks dim enough that many would do things like increase the gain or ISO on their cameras to boost it up and make it brighter. But, the truth is, it's just as good an image as the bright 10% noise case above.

Here is that same image, stretched to restore the image intensity (multiplied by 2).



If you compare this to the original version with 10% noise, you'll be hard-pressed to tell them apart. Sure, the precise pattern of noise here is different, but to the eye (and according to the math), they've got the same SNR. The intensity in that region is now 85 and the standard deviation is now 25.8 for an SNR of 3.29. (Note: Don't be bothered by the slightly different numbers here like 3.27 vs. 3.29. These are just rounding errors.)

***Note:***Here, I've been calculating SNR just by that simple ratio. You'll often see it expressed in dB (decibels). To calculate things in dB, just take 20 \* log10 (ratio). So, if we had 2,000 for our signal and 2 for our noise we'd have 20\*log10 (2000/2) or 60 dB.

*Upshot: SNR*

The upshot here is to remember that SNR has two components: the signal and the noise. If we keep the signal constant and halve the noise, we double our SNR. If we keep the noise constant and double the signal, we double the SNR. If we halve the signal but cut the noise to a quarter of what it was, we also double the SNR. It's a ratio and getting better SNR can come down to either boosting the signal or dropping the noise.

*Types of Noise*

**Read Noise**

Every time you read an image off the camera, some noise is going to be injected into the image. Even if there is no signal (no light, no appreciable dark current), you're still going to have noise in the image (it won't be perfectly even). This noise is called ***read noise*** and it comes from both the electronics on the sensor itself and from the electronics inside your camera.

Read noise comes in several flavours. Ideally, the noise has no structure to it. That is, there is no fixed pattern to it and it just an even level of background, Gaussian noise. If you were to imagine this as sound, it would be a simple hiss without any tones, clicks, pops, etc. Visually, this kind of noise is easy to overlook and when images with simple, Gaussian noise are stacked, the noise level is reduced (by a factor of the square root of the number of stacked images).

Many cameras aren't ideal, however. Some have a fixed, spatial pattern to the noise. Every time you take an image, there is a clear pattern of streaks, waves, etc. in the image. If you're going to have something sub-optimal, this is what you'd like to have as it's entirely repeatable. Yes, it will be in every light frame you take, but it will also be in every dark frame or bias frame you take, so our standard pre-processing steps can remove it.

Other cameras have noise that is at a certain frequency in the image (e.g., vertical streaks that are 100 pixels apart), but with the position varying from image to image. This is a lot tougher to deal with as each frame will have the noise pattern in it, but no two frames will be alike. So, the same pixels won't be affected in your light frames and your dark frames, making standard pre-processing ineffective (it won't subtract out). With this kind of noise, all you can really do is either try to filter it out later or try to stack enough frames so that the noise will end up being even across the final image (each frame injects the noise somewhere else). Unfortunately, that can often require a lot of frames to do. Worse still, since your darks and biases have these same issues, you will need a lot of darks or biases to make that pattern disappear. If it doesn't you'll end up injecting noise into your lights when you subtract out your darks or biases.

Finally, one other kind of noise should be touched upon here. Salt-and-pepper noise looks like bright and dark spots that are scattered across your image. These spots tend to vary from image to image (we're not talking about hot pixels here). Some cameras have real issues with these and if your camera does, you'll need to take measures to remove them. Stacking based on the standard deviation across images (aka *sigma-clip*) can be very helpful.

**Shot Noise**

If you're a doctor on the battlegrounds of Prussia in the late 1800's and wondering when you'll see the next person come in having been kicked by a horse you have something in common with an astrophotographer wondering (albeit very quickly) when the next photon from that galaxy will arrive at the sensor. Both of you are thinking about Poisson processes. Well, you're probably not overtly doing this, but a guy named Bortkiewicz was. Sometimes he'd see one person arrive in the morning and another that afternoon and other times he'd go days without seeing any. The point is that there is some likelihood of an event (getting kicked by a horse or having a photon from a galaxy make it to your sensor) and since it's just a probability and not a certainty, there will be some variation in how long it is between events. That variation was described back in the early 1800's by the French mathematician Poisson.

If you take an image for a second, you might get 10 photons from that DSO to hit a pixel on your sensor. Take another and you might get 12. Take another and you might get 9. Because of the way light is, we have this variability. This variability is called ***shot noise*** (also called *photon*noise) and it follows that Poisson distribution.

You're never going to escape shot noise and the amount of noise you have goes up only by the square root of the intensity. So, the brighter the object, the more noise you have. That sounds bad, but in truth, it's really not a problem as the signal went up as well. So, if you have 100 photons hitting your sensor in a shot in one case and 10 photons hitting in another case, your SNR is over 3x higher with the greater number of photons. Despite having higher noise, the higher signal more than overcame this (SNR would be N/sqrt(N)).

So we can ignore shot noise, right? Wrong. There are two sources of shot noise. We don't mind the shot noise from the DSO itself, but what about the shot noise from the skyglow? Your sensor doesn't know if the photons are streaming in from a galaxy or from the skyglow. Both have shot noise. This is what is so evil about skyglow. If it just brightened the image, all we would need to do is to slide our black point up and cut off the skyglow. But, ***skyglow lowers the SNR by injecting shot noise into the image without also injecting signal into the image***. That's one of the key reasons (that and not lowering the effective dynamic range of your chip) why skyglow is so harmful.

**Dark Noise**

Block all light from hitting your sensor and take images of varying exposure durations and you'll notice that the image gets brighter with increasing exposure. This results from ***dark*current.**The intensity should double as you double the exposure duration and it should also double for every 6 degrees Centigrade or so. You'll also find that some pixels brighten faster than others (hot pixels brighten very quickly), leading to a **pattern** of fixed, spatial noise. This is why we typically take dark frames (using the same exposure duration and temperature so that the **pattern** in our darks is the same as the **pattern** in the lights - meaning, we can subtract one from the other to arrive at a cleaner light frame.)

Most all of you reading this will know about dark frames, but some number of you may not have considered one real implication of dark current. Remember shot noise? Since photons arrive in a random process, we don't really know exactly how many we'll get and that the variance in the actual count is proportional to the signal level? The same holds true for dark current. The higher the dark current, the more variability there is in our reading in an individual frame. Therefore, if we want a very good estimate of the average dark current (so we can subtract this *expected value* of the dark current from our lights), we need even more dark frames to average together.

Thus, the answer to the question, "How many dark frames should I use?" depends on your dark current. If you've got a lot of it and your image gets considerably brighter and noisier as you take longer dark frames, you're going to need more darks. If you don't collect enough darks, you're going to inject noise into the light frames when you pre-process the images. Any deviation in your dark stack from that expected value of the dark current for each pixel means you'll be injecting that difference into each of your lights***. My note- so the pattern noise element can be removed but not the associated random shot noise. Turning off ‘Live View’, cooling the camera and letting it cool down between exposures can help reduce heat production inside the camera.***

**Quantization Error**

When we read the collected electrons (as a voltage) off of a sensor, it's an analog signal. This is turned into a number using an analog-digital-converter (ADC) on the sensor. Suppose you have an 8-bit ADC. With this setup, you can record 256 shades of intensity (2 raised to the 8th power). Now, suppose further that your CCD can store 10,000 electrons before it saturates. If you want to use the full dynamic range of the CCD, you must setup the ADC so that each integer in the scale here represents about 25 photons. So, a 10 coming off the ADC would mean that there were about 250 photons captured and an 11 would mean there were about 275.

It doesn't take much to see the problem here. You can no longer tell the difference between 250 photons and 251 or 255 or 260 photons. They all get the same value. That problem is called ***quantization error*** and it turns similar, but not identical intensity values (e.g., the nice subtle differences in shading in that galaxy arm you hope to pull apart) into the same value. Once it's the same, information is lost and there's no pulling it apart (at least not in one frame).

Quantization error comes about when you've got more levels of intensity to store than you've got numbers to store them in. These days, dedicated astro-cams don't suffer from quantization error, but DSLRs and lunar / planetary cams still can. For example, if you've got a CCD that has a full-well capacity of 10,000 electrons and you've got a 12-bit ADC that gives you 4,096 intensity values, you've potentially got a problem. Here, the ***system gain*** would have to be about 2.5 e-/ADU if you wanted to use the full range of the CCD. That is, each intensity step (Analog Digital Unit) would represent 2.5 electrons. Alternatively, you could potentially set the system gain to 1 e-/ADU and trade off ***dynamic range*** to get you out of the quantization error. You'd saturate the ADC at 4,096 electrons, but you'd not have any quantization error to worry about. Cameras that let you vary the system gain (e.g., DSLRs - they call this ISO) let you play both sides of this trade-off.

Before leaving quantization error, I should note one more thing. We should not be overly harsh on cameras that have a 12-bit ADC and that have full-well capacities greater than 4,096 electrons. In truth, you would be hard-pressed to see any difference between a sensor with 10k e- of full-well capacity hooked up to a 12-bit ADC from one hooked up to a 16-bit ADC (which has 65,536 intensity steps). Why? The noise. The noise is already limiting what you can resolve. Let's say the 10,000 e- sensor has 10 e- worth of noise on it. The max SNR (aka the dynamic range) there is 1000 (aka 60 dB) as we know the last digit there is really run by the noise. Our 12-bit ADC with 4,096 intensity values (aka 72 dB) now doesn't look so bad. Run that same 10,000 e- sensor though with only 1 e- worth of noise and we have a dynamic range of 10,000 (aka 80 dB) and we're back to having a problem. The idea here is just because you have some # of electrons worth of full-well capacity doesn't mean you can really know exactly how many are in any given well (the noise) and that uncertainty means you don't have to have discrete values in your ADC for each possible option. Personally, I'd not worry a minute if my 16-bit ADC (65,536 steps) were hooked to a CCD with 100k e- worth of full-well.

*Upshot: Noise*

You're going to have noise. No matter what, some will be there. The trick is, of course, to minimize it. Much of the noise will come from inside the camera and the choice of camera is going to factor in a lot here. You want something with noise that is both low in level and well-behaved. We'll cover how to measure this in Part 2. Other noise will come from what you're shooting-both the target and the sky. The target's **(random)** noise you can ignore (nothing you'll ever do about it), but the sky's noise is something that you may be able to work against. We'll cover this more in Part 3.

*Signal*

Boosting SNR can come from dropping the noise or boosting the signal. So, it's worth considering where the signal comes from and how we might go about boosting it. The first answer as to where it comes from is an obvious one. It comes from that faint fuzzy you're trying to image (duh!). Moving beyond this, we realize it must get from that faint fuzzy to the pixel on your sensor. Along the way, there are a few things to consider that affect the amount of signal you're getting.

The first is the **aperture** of your scope. The bigger it is, the more photons it sucks in and directs towards your sensor. A variant on this, though, is the **f-ratio** you're working at. There has been a lot written on this on the Internet and this is something I took on in a [blog post of mine](http://www.stark-labs.com/blog/files/FratioAperture.php) and will cover here eventually. The f-ratio, though, really is determining how much signal is hitting a given pixel when one is considering extended objects. If you keep the aperture constant and vary the f-ratio (by varying the focal length), you're **trading off signal and resolution**. Long focal lengths spread the image over more pixels giving you higher resolution, but cut the light hitting each pixel. The point of the "[f-ratio myth](http://www.stanmooreastro.com/f_ratio_myth.htm)" argument is that **once we're above the read noise, the added resolution comes at little or no cost.** Running longer focal lengths (higher f-ratios) gets you closer to this noise floor, however. But, the details of this are for another day. What's important here is that if you want to boost the SNR in a given pixel (and yes, you may not want/need to if you're already bright enough, but if you do ...) you can do so by dropping the f-ratio (increasing the aperture for a given focal length or decreasing the focal length for a given aperture).

The second thing to consider is related to this and it's the ***light throughput*** of your scope. Let's consider a Newtonian with simple Al coatings that give you about 87% light reflectivity have two surfaces, each at 87% making for a total throughput of 0.87 \* 0.87 or about 76%. If we call this an 8" f/5 mirror with a 2.5" obstruction, our total throughput drops down to 66% of the light that came into the tube. This means we're working like an ideal 6.4" scope in terms of our light gathering. Now, let's boost those coatings to 95%. Before the obstruction, we're running 90% and with that in place, we're running 80% efficiency and like a perfect 7.1" scope. Hmmm... gettting those old mirrors recoated with something fancy just gave us a 14% boost in pure signal.

The third thing to consider is the ***quantum efficiency*** (QE) of your camera. QE refers to the proportion of photons that are recorded out of the total number that hit the sensor. If you've got one camera with a QE of 40% and another with a QE of 80% and all else is equal, the one with the higher QE has twice the SNR as the one with the lower QE. Another way to think of it is that the boost in QE is like a boost in the aperture of your scope. Run a QE=80% camera on a 80 mm APO and you'll do as well as a QE=40% camera on a 113 mm APO (focal length and noise figures being equal). Yet one more way to think of it is that if you want to capture some number of photons to get your signal above the noise, it will take twice as much imaging time on the lower QE sensor.

One thing not to consider here is the darkness of the skies. People often think of this as a boost in the signal as the DSO sure seems brighter when you look at it. But, how would M51's photons know that you're in an urban vs. rural area? They don't, of course. The same number of photons are streaming towards you from M51 and it's not like the photons shooting up in the sky from the city lights are deflecting them. No, the signal is the same and it's just that the background level has been raised making it tougher to pull that signal out.

*It's all about SNR*

Remember at the outset when I said all those things come down to SNR? Aperture rules, right? Yes it does and it does because it boosts the SNR. Whether attached to a 3" or a 30" scope, your camera has the same read noise and the same thermal noise. Attached to a 30" scope, however, your camera is getting hit with 100x as many photons as when it's attached to a 3" scope (for the sake of the argument here, assume equal focal lengths.) That's quite a boost to the signal.

OK, that was easy. What about dark skies vs. urban skies? Dark skies have greater contrast and let you shoot longer, right? How's that SNR? Two things come in to play here. First, we can consider the case in which you expose for the same duration under both conditions. Here, the read noise and thermal noise are the same and the number of photons from the DSO are the same. So the signal and two components of the noise are the same. But, the shot noise is different. The camera doesn't know skyglow from the DSO as photons from the sky are photons from the sky. In our urban site, we may have 4 times as many photons hitting our sensor from skyglow as we do in our city site. While we can stretch and set a black point to make the sky dark again, all those extra photons had all that extra shot noise. In fact, 4x the skyglow will lead to a doubling of the shot noise. Under urban skies, shot noise from the skyglow is usually the dominant source of noise, swamping out the read noise and thermal noise. So, we've kept the signal but doubled the noise and halved our SNR. **Own note-Samir Kharusi has calcualted the difference in skyfog can be x10-x40.**

Monochrome vs. color cameras? Color filters cut your signal down to under a third of what it would have been without the color filters. You've cut the signal and kept the noise the same. Camera choice? Boost the QE and you boost the signal. Drop the noise and you've, well, dropped the noise. Either way, you've boosted the SNR. **Own note-removing the internal filters in the DSLR improves QE in red channel and slightly in the Blue and Green.**

*What's Next?*

Hopefully, most of the entries in this column won't be quite so long as this one. But, there was a lot of information to get through at the outset here. Coming up in Part 2 will be a tutorial on measuring the noise in your camera. It's really a lot easier than it sounds. After that, Part 3 will cover things you can do to try to improve the SNR in your images and how thinking about SNR may change the way you take your shots.

**Signal to Noise: Understanding it, Measuring it, and Improving it | Part 2 - Understanding One Pixel**

May 26 2009

In [Part 1 of of this series](http://www.cloudynights.com/item.php?item_id=1966), we went over the basic notions of what we mean by signal and what we mean by noise. Our images will always have noise and we're never going to get around that. The trick to making our pictures look pretty is to have enough signal to noise ratio (SNR) so that we can stretch well enough to resolve those faint bits without having the image look like a mess. As SNR on the bright bits is easy, we will spend most of our time talking about the low end of the intensity scale.   
  
In this instalment, we'll talk about the fundamental unit of our image – the pixel. An understanding of the pixel is the fundamental to understanding our image and data quality.   
  
Before I get going, I do want to mention that it's not like I'm blazing a lot of new ground here. There have been a number of attempts to explain SNR in the pixel on the web. One well-known source is [Steve Cannistra's coverage](http://www.starrywonders.com/snr.html) and another here on Cloudy Nights is [Charles Anstey's](http://www.cloudynights.com/item.php?item_id=1622). I've also recently come across [Sam Fahmie's excellent article](http://www.samfahmie.com/astronoise/noise.html) on the topic. Each of these covers the math in great detail and these are great resources. What I hope to do here is to be faithful to the math (numbers don't lie – people lie with numbers!), but to try to cache this in a different format. I'll use a few equations, but nothing will get tougher than what the simple calculator program on your computer can handle. I'll toss numbers around (quite a bit actually), but the goal of this is to show you where they come from and how you can get to them yourself. To help out, I've also [provided a spreadsheet](http://www.stark-labs.com/craig/articles/assets/SNR%20Calculator.xls) that you can use in Excel, Open Office, Neo Office, or anything that can work with an Excel spreadsheet.

As we work through this and through the upcoming parts of this series, you may note a particular philosophy of mine. If you don't, I'll lay it out plainly here. It may seem odd coming from someone who makes a living as a scientist and researcher and someone who tries to do astrophotography, but the philosophy is to ***not stress and get hung up on being incredibly precise about this***. I've got a good buddy, we'll call him Michael (because that's his name), who was running some of the calculations to figure the optimal sub-exposure duration and was coming up with numbers like 3 minutes and 17 seconds. He had a nice library of 3 minute darks and was about to go and make a new one at 3 minutes and 17 seconds before I talked some sense into him. For starters, he'll never see the difference even if the extra 17 seconds puts him in a theoretically better place as the difference is trivial. Second, the calculations are all based on estimates of values and, if these estimates are a bit off, so are the calculations. So, don't sweat the details too much here. So, if we end up at a place in a future article saying that 1”/pixel is about as high a resolution as you should go and you're running 1.1” or 0.9”, don't go out and swap cameras, scopes, reducers, etc. to get to exactly 1”. Those differences are trivial and the suggestion is based on estimates of seeing, tracking, etc. But, if you're at 0.1” per pixel...it becomes a good time to come up with a better solution for astro-imaging. 

**Sources of Signal and Noise**

Photons are hitting your CCD/CMOS – it really doesn't matter here and the sensors job is to count those photons. You get an analog measure of the photon count out of each CCD well since it is made up from the electrons stored there while imaging and this gets passed to an analog-digital-converter (ADC) turning that analog measure into a number. That number is a scaled version of the number of photons that hit the CCD well, plus or minus something. That plus or minus value is the noise or variability in our estimate of the true number of photons that have hit the pixel.   
  
As discussed last time, there are a few sources of signals and noise to contend with here. Let's briefly recap them in terms of what goes into that CCD well:

1. We have our photons from our faint fuzzy. For example, we may have 10 photons per second coming from the DSO. These are good. We like them. Let's call them the real signal.
2. We have photons coming from the sky but not from our DSO. These are skyglow-based photons. Perhaps we have 100 photons coming from skyglow into our image per second. So far, skyglow isn't really all that bad (but it will be in just a second). It's signal, but signal that we don't really care about.
3. We have photons coming from the dark. OK, they're not coming from the dark *per se*, but even when it's totally dark we have these getting picked up by our CCD well at a certain rate. The rate will vary a bit from pixel to pixel (and some pixels have a very high rate – hot pixels). But, if we consider just one pixel at a time, this rate is constant. Let's call it 0.03 per second for a cooled camera. Again, it's signal (it made the image brighter), but signal we'd just clip off in processing.

For what it's worth, these numbers aren't entirely made up. I just looked at a 1-minute image of M51 taken from my urban skies on a QSI 540 and a Borg 101 ED f/4. The background was ~4200 ADU, a spot of the arm was at ~4500 ADU, the average dark frame was ~215 ADU, and the average bias frame was ~214 ADU. The gain here is 0.8 e-/ADU and peak QE is 55% with 50% or so being a decent average value for the range recorded. So, if we remove the bias, that's ~4000 ADU of background, ~4300 of background plus target (~300 then of just the target when we remove the background as well), and ~1 ADU of dark current. Convert this into photons (multiply by 0.8 to get into electrons and then by 2 to get into photons that hit the detector, only half of which were recorded) we get 6400 photons per minute from skyglow, 480 per minute from my target, and 1.6 per minute from dark current (that is 107 photons per second, 8 photons per second, and 0.03 photons per second). Now this isn't it. If it were, we'd be dancing. But we also have:

1. Shot noise from the target. Photons don't arrive at a nice, constant rate. So, there is an uncertainty here of sqrt(N) in our signal (where N is the photon count). In our example, in one second this is about 3.2 photons of variability (noise).
2. Shot noise from the skyglow.
3. Shot noise from the dark current.
4. Read noise.

**The Pixel SNR Equation**   
  
Since SNR is a ratio (signal to noise ratio), we've got two parts to consider. First, we've got the “signal”. That's got the first three parts laid out above. The real signal we want from our DSO, the skyglow, and the dark current.

*Total\_Signal = Duration \* (Target + Skyglow + Dark)*

We can think of this in terms of photons or electrons (recall however only about half my photons were captured) and here, we can even think of it in ADU. All this says is we've got three sources of "flux" (streams of photons coming into our detector at certain rates). Add them up you've got a total rate per second. Multiply that by your exposure duration and you've got the total signal. No rocket science here.

*Shot noise = sqrt (Total\_Signal)*

Recall from above and from [Part 1](http://www.cloudynights.com/item.php?item_id=1966) that we've got this uncertainty in the actual photon count that goes along with the the square root of the expected number. So, we've got that as part of our noise.

*Read noise = ... well ... read noise*

Every time you read off an image, your camera will inject some number of electrons of read noise.   
  
Now, we need to combine these two sources of noise. We do that by taking the square root of each of them squared. So, we have sqrt(sqrt(Total\_Signal)2 + Read\_Noise2) which simplifies to:

*Noise = sqrt(Total\_Signal + Read\_Noise2)*

For now, this is all the math we need. Let's stick some numbers in here and let's work in electrons instead of photons. Based on what I recorded from that image of M51 we have:

|  |  |
| --- | --- |
| **Signal source** | **e-/sec** |
| Target | 5 |
| Skyglow | 50 |
| Dark current | 0.02 |

This makes our equation for the signal be:

*Total\_Signal = Duration \* 55.02*

If we plug in 60 seconds here, we get a total signal of ~3300 electrons. (Recall my camera has a gain of 0.8 e-/ADU so if we divide this by 0.8 we get a bit north of 4000 ADU, nicely back where we started give or take the rounding I did for nice, easy numbers).   
  
Now, a typical value for a good camera is about 8 e- of read noise. If we figure our noise in this 60 second exposure we get:

*Noise = sqrt(3300 + 82) = sqrt(3364) = 58*

At this point, you might be tempted to just say the SNR is 3300 / 58 or about 57 and be dancing as that's a very good SNR for a single frame. But, remember that a good bunch of this “signal” is signal we don't care about. We're going to take our Levels tool or our B slider and "reset the black point" to remove all that skyglow (and dark current). We're throwing that away as we don't care to know just how nicely bright our skies are. Our real signal is the signal from the DSO. That is 60 seconds \* 5 e-/second or 300 electrons. Our SNR is now 5.2. Better than an SNR of 1 but it's not 57.   
  
   
  
Here's the shot (raw, no pre-processing, just simple linear stretching) you can see what this level of SNR looks like. You can see M51's arms above the noise. Our total SNR will therefore have to be a good deal better than this in the end. 

Example-  
  
   
  
The data from that M51 shot could leave you thinking that dark current and read noise don't mean diddly (technical term). The read noise was 8 e- and the shot noise from the dark current was nil. The shot noise from sky glow was almost 55 e-. Since the total noise was 58 e- (remember, the way to add noise isn't just to add them but to square each, sum, then take the square root), that 55 e- of shot noise from the sky is pretty daunting. Let's consider another case though and look at some data shot through an H-a filter using the same camera and scope. Here, the background sky near the Cone Nebula reads about 440 ADU in a 10 minute exposure and the nebula itself is up around 505 ADU. Dark frames are at about 216 ADU still. If we pull the 214 ADU bias signal out we're left with 2 ADU from the dark current (1.6 e-), 224 ADU from skyglow (179 e-) and 65 ADU (52 e-) from a spot on the target.   
  
So, the signal we care about is 52 e- and the Total\_Signal is 233 e- (target + skyglow + dark). Our noise is then:

*sqrt(233 + 82) = 17.2*

The SNR in this image is 3.0 (a bit worse than before, but much of this depends on where you look in the image). The real point here, though is to consider where the noise is coming from. Overall, there is ***a lot less noise here than in the earlier case.*** In fact you can see it in the image. Take a look at the background of the Cone here and of M51. Notice all the extra noise in the M51 background? That noise is from the skyglow's shot noise.   
In the Cone shot here, we have 17.2 e- of noise and before we had 58 e- of noise. Last time, if we ignored read noise, we had 57 e- of total noise. Removing the shot noise from the target, we still had 55 e- of noise just from the skyglow. This, out of a total of 58 e- of noise. In this case, who cares about the read noise? Our SNR considering the read noise was 5.2. Pull it out and the SNR is 5.3.   
  
Here, though, we have a different story. Here, our total noise term is 17.2. If we just look at the shot noise overall it's 15.2 e-. Shot noise from the sky is 13.4 e- and the read noise is 8 e-. These are a lot closer. SNR with the read noise in there is 3 and pulling it out, it's 3.4. Clearly, it's more of a factor now. If we could eliminate it, our SNR would go up 13% whereas previously it would go up under 2%.   
  
**This is why you may have heard people say that, oddly enough, a pristine camera, low read noise, low dark current, etc. are more important at dark sky sites.** This is also why the various calculators for optimal exposure durations say you should use longer exposures at dark sky sites than under urban skies. You might think that the urban skies present more of a challenge (they do) and therefore need a cleaner camera. **The truth is that the urban skies themselves are giving you so much noise that you'll be hard-pressed to take advantage of the lower noise offered by better cameras.** (Rather large caveat – all this assumes that, when comparing cameras, both have “well-behaved” noise. Many don't and this will be covered in a later article). 

**Playing with the numbers**

At this point, I hope you're sitting there scratching your head and saying “Hmmm...”. Perhaps there are questions like: What kind of SNR do I have in that galaxy arm? What kinds of SNRs give me shots I like? What would happen if I moved your scope to a darker site? Is that filter really helping me? Does my buddy who claims his light pollution is worse than mine have a leg to stand on or is his SNR no worse than mine? What would happen if I got a camera with lower dark current?   
  
Playing with the numbers lets you get at these issues. I call it playing because it should be fun. Again, don't stress over the details. If you sample a spot in the image and see it's at 1017 ADU in one place but 1024 right next door, don't sweat it. 1020 is a fine number. The goal is to get a handle for how some of these things behave and to see where the big wins and losses are and what may not matter. I've put together a spreadsheet in Excel format to let you start to get a handle on this. You don't actually need Excel and, in fact, I didn't put it together in Excel. OpenOffice is a perfectly fine stand-in (and is free) and Google Docs can do this as well.

**Spreadsheet basics**

To use the spreadsheet, start by finding a light, a matched dark, and a bias. Open up the light and you'll be looking for two values. One is just the intensity of the skyglow background and the other is the intensity of the target (pick some area that's not the galaxy's core, but rather is that arm or bit of nebulosity you want to pull out). These should both be from a raw light frame (no pre-processing). Enter in the intensity of the dark frame and of a bias frame and enter in the exposure duration for your light and your dark (make sure your dark frame is at least as bright as your bias frame). Finally, enter two parameters about your camera that you should be able to get from the camera maker (or any of my reviews): the gain and the read noise.   
  
The bottom half of the first sheet shows the calculations I've performed for the samples here. By default, it loads up with sample data from my urban back yard. You'll see that the SNR of the part of the arm in M51 I sampled is about 4 (the numbers used in the example above were rounded a bit, but the spreadsheet has the actual values from my yard). Supplied in there are also values from the L channel of a [shot of the Veil nebula I took](http://www.stark-labs.com/craig/photos/photos.html). This was out in Julian at my buddy Chuck Kimball's [“Big Cat Cabin” B&B.](http://www.artistsloft.com/) He's got much darker skies as you'll see. It'd take 20 minutes of exposure to reach the same level of background glow out there that it took me to hit in only 1 minute! **My note-remember the calculation by Samir Kharusi that sky fog can be x10-x40 greater in an urban area than a dark site.**

**Fun with graphs**

The fun doesn't end with the simple calculations. Included in the spreadsheet are several other sheets with graphs to show various “what if” scenarios. The first extra sheet (Skyglow effect) shows what would happen to your single-frame SNR if you changed to a brighter or darker sky site. Sky glow has a pretty bad effect on SNR.   
  
The next (*Duration effect*) shows the effect of exposure duration on SNR in a single frame. Certainly, the more data you get the better as the SNR goes up with exposure duration. Recall that read noise is always there and the skyglow's shot noise goes up with the square root of the duration. The target's signal goes up with duration so that's why the single-frame SNR continues to rise. Of course, you could have taken four shots at 5 minutes and stacked them rather than one at 20 minutes. That effect is plotted in another sheet (*Sub-exposure length*).

Next is one called *Stacking*. It shows what will happen to your SNR as you take multiple exposures (at your current settings) and stack them (again, assuming well-behaved noise on your camera's part). Here, you'll see just why we stack! You can also, though, do things like look at the SNR of a single 10 minute exposure on the *Duration effect* sheet and compare it to 10 frames at one minute here (or use whatever works out given the exposure time from your own data. 

Finally, the last one, called *Sub-exposure length* lets you enter a fixed total imaging time (default is 180 minutes) and examine the SNR in a stack as a function of the duration of individual sub-exposures.

Note, in each of these, there are a number of simplifications. Most notable is the fact that we're assuming either no dark subtraction or infinite SNR on your darks. Likewise, your flats are assumed to have no noise. Each of these steps introduces noise into your image in the process of correcting artifacts. Despite these limitations, they can give you a good feel for how things affect your SNR.

**Sample “What if?” Experiment: Effect of Skyglow** 

Previously I gave the example of the Veil shot at Chuck's site in Julian and my yard here. We can use this sheet to see just what would happen if I tried this same M51 shot but out in Julian. This serves as a sample of the kind of thing you can do when you start playing with the numbers here and it lets you see just what you've traded off.   
  
If we use my original parameters, we see that my sky glow flux is 53 e-/second and the target flux there on M51's arm was 4 e-/second. Let's transport my rig to Chuck's site where the sky glow flux is only 2.36 e-/sec. We can't just paste the parameters from the Veil in there and be done as that's the Veil and we're shooting M51. So, we need to twiddle a bit. If we first adjust the sky glow background down to 390, we get close to Chuck's sky glow flux. Of course, the target + sky glow won't be up at 4500 where we started. We want 300 ADU of target signal (where we were before) which should lead to the target flux of 4 e-/sec (it's not like M51 is any brighter out there than here – it's just that his skies are darker than mine). So, if we set the target signal in the top there to 690 (390 + 300) we get back to the right flux from M51.   
  
Take a look at the SNR how here. You'll see we have an SNR of 11.38 in our section of M51's arm there from the dark site. Recall that this same spot had an SNR of 4.06 from my urban yard. We're talking almost 3x the SNR here and recall that you need 4x the data to get 2x the SNR (1/sqrt(N)). Paste back in the original values and then head on over to the *Stacking* sheet and you'll see it takes about 8 stacked frames to get to this SNR (look at the graph or the table on the right). So, this says we need 8 times as much data from the urban site to equal the dark sky site.   
  
Paste back in the dark-sky values and head over to the *Sub-exposure length* sheet and you'll see now that we're over 160 on our SNR with most sub-exposure lengths. That's what 3 hours at Chuck's site will do. Now, what kind of total exposure time do I need from home to hit this? Paste back in my yard's parameters. In 3 hours, we're below 55 on the SNR. Now, up the *Total duration* field here. Start upping my total exposure time and look at the effect. You'll need to get to about 1550 before my SNR matches Chuck's cabins. That's about 26 hours. Ouch! When doing LRGB work, I really should get out there to his B&B (or to my club's dark sky site). (Note, how 26 hours is about 8x the 3 hours and recall the 8x factor we came up with just above?) Armed with this we can now start to ask other questions such as what about what if I took images but shot instead only with a Ha filter in the optical train? Get the idea? 

**Conclusions**   
  
Coming down the pike are going to be articles on things like sampling and how you can view this as a trade-off between SNR and resolution. Along for the ride in that one will be a discussion of f-ratio and just what it does and doesn't mean to us. Color vs. mono cameras will be covered as well. We'll go a bit deeper into things like how pre-processing can add noise into your images as well. Each of these requires a handle on SNR in one pixel. Along the way, we'll also cover how to measure things like read noise or system gain in your camera without any fancy equipment. Hopefully, the first two articles have filled in any gaps in your knowledge here.

A few things to note before closing. First, is that one of the great things about [Cloudy Nights](http://www.cloudynights.com/) is that we've got such great discussion boards. There are threads there for each of the articles and I'm reading them. **If something doesn't make sense, speak up!** I'm a professor by day and try to convince my students that if they've done their job by coming to class prepared and paying attention and still don't get something, it's my fault not theirs. The same kind of logic applies here even though the roles aren't exactly the same. If you're reading along and trying to get it but things still don't compute, odds are I skipped a step and/or wasn't as clear as I could have been. If this is indeed the case for you then please do speak up on CN..   
  
Second, Mike and Mike here have given the A-OK for me to post PDFs of these up on [my personal website](http://www.stark-labs.com/craig/articles/articles.html). So, if you're looking to print out hard copies for some offline reading (I won't ask where), that's a good place to start.

# Signal to Noise: Part 3 - Measuring your Camera

Jul 06 2009

|  |
| --- |
| In [Part 1](http://www.cloudynights.com/item.php?item_id=1966), we covered the basic notion of SNR and in [Part 2](http://www.cloudynights.com/item.php?item_id=1973), we covered SNR in a single pixel. If you've not read those bits yet, head back and give them a look. If you have, you've had your nose in the books for a bit now and it's time for a break. In this instalment, we're going to take on the practical aspect of testing your camera and figuring out just what kinds of camera noise you're up against. Warning - this is a long one. You may find it helpful to grab the PDF of the article that I've put up (along with others in the series) on [my personal website](http://www.stark-labs.com/craig/articles/articles.html).  Believe it or not, you can get very accurate measurements on your camera with only a minimum of hardware, skills, and time. Here, you will set out to measure:   * System gain (number of electrons per ADU) * Read noise (in ADU and electrons) * Dark current (in ADU and electrons) * Dark current stability   You can also go on to probe some of the inner workings of your camera and look for it's "fingerprint" by some detailed analyses of the read noise. Here, we'll look at:   * Histogram of the read-noise (just how Gaussian is it?) * Amount of fixed-frequency / variable-location noise (the worst kind!)   Believe it or not, you only need a few tools to do all of this and all of this can be done with the camera sitting on the desk next to you (no need for the telescope). All we need is the ability to take clean dark frames and to take reasonable flat frames. So, here's a parts-list:   * (optional) An SLR camera lens you can attach to your camera. If you've got this, you can take better flats and control the amount of light hitting your chip. If not, you'll live. * A metal lenscap for the camera or SLR lens. If you don't happen to have this, a piece of tin foil and a rubber band will do. * About 15 sheets of white paper roughly 4x4 inches or so apiece. * [ImageJ](http://rsbweb.nih.gov/ij/). Freeware image processing and analysis software. Much can be done in other programs, but ImageJ does give you nice FFTs * (optional) A spreadsheet program to graph your results and do things like fit a linear regression line.  *Getting the Data* We're going to collect a bunch of bias frames, several dark frames, and some flat frames. Get your camera setup, but don't get it turned on and going yet. Keep it at ambient temperature. I do all of this on a desk without a telescope attached, as there is no need at this point for any kind of lens.  First up are the biases and darks. Here, we need to make sure that no light whatsoever is getting to the sensor. Believe it or not, black plastic lenscaps are often pretty transparent to IR light. This is why I said you need a metal lenscap. A perfectly good solution is to use your camera's 1.25" or 2" nosepiece and to wrap a piece of tin foil over the nosepiece. Hold it all in place with a rubber band. Voila! Perfect dark frames.  Most cameras these days are very light-tight, but some still aren't. If you're worried that yours isn't (or you know it isn't -- look at a dark frame and see if there is one side that's brighter than another), you'll need to shade the camera body from any ambient light. One way to do this is to work in the dark (just don't aim your computer screen at the camera). Another is to put a box over the camera. If you go that route, make sure there's enough ventilation still to keep the camera from getting abnormally hot. Your goal here is typically to keep the camera shaded and not in direct light. If it can't deal with a small amount of reflected light, you've got bigger fish to fry.  Now, fire up the camera and connect in your capture software. Right away, fire off:   1. A set of 1-minute dark frames (at least 30). These are used for your dark stability measurement. Since the camera is at ambient, we get to see how its dark current changes as you get going. If you've got cooling, it'll start to drop as you head towards your set-point or the max-cool level. If not, your camera will start to warm up here. Once done, your camera should be at some kind of thermal equilibrium. Some cameras do need more than 30 minutes to hit this, though. 2. A large stack of bias frames. These will be used in a number of measures. These days, I grab 250 of them to be safe. 50 would probably do just fine, but unless there's a compelling reason not to, grab at least 100. The exposure duration here is typically set to 1 ms, so it's not like this should take a long time. Note, even if your camera isn't thermally equilibrated yet, you can still get these. At 1 ms, there's no dark current to speak of. 3. A set of dark frames of varying length. These will be used in calculating the dark current. I typically grab 1 m, 2 m, 5 m, and 10 m frames. Note, if you think your camera may have thermal stability issues (i.e., it's uncooled), you may want to space these out. So, wait 15 min or so after the bias frames and then grab the 1 m frame. Wait 15 min again and get the 2 m. The goal here is to let the camera's temperature stabilize.   After this, you'll want to setup for flats. This will enable you to calculate the system gain of your camera. Here, you need to have the ability to take pairs of flat frames of varying brightnesses, ideally without changing the exposure duration. The goal here is to take pairs of flats, perfectly matched in intensity, for a range of intensities. There are two methods that I've used:   * An EL-panel with variable brightness that sits on the front of an SLR lens (camera and lens aimed up). This is the uber-cool way to grab the flats as you can dial in any brightness desired. Set the exposure duration to something like 0.1 s and the panel to something dim and grab a test shot. Adjust the exposure duration and/or f-ratio on the lens to be above the level of a bias frame by just a hair. You can now adjust the brightness of the panel to increase the brightness of the flat. * A stack of white office paper acting as a diffuser. Start with ~4 sheets on the nose of the camera or on the front of your lens and aim the rig at something like a white ceiling. You won't need to be perfectly flat and this will get you quite close. Setup for a short duration (e.g. 0.1 s or so) and use your capture program's histogram to see how bright the image really is. The goal here is to be near but not entirely at the top of the histogram. To adjust the brightness of the flat, you'll simply add another piece of paper onto the stack.   https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m5b9c7f8f.pngNow, take *pairs of flats* at various brightness levels. Make sure that the overall brightness level covers a good range of the intensity scale. The figure here shows the histograms (plotted from *Nebulosity*) for ten different intensity levels I used in testing the Atik 314L+ here. You don't want to bottom out and be looking like a bias frame, but you don't want to saturate the CCD either. Err on the side of being in the lower-half here as above this, your sensor may be non-linear. I'll typically use at least five brightness levels. You can do this with just one, but you'll be less prone to error with more. Make sure you name them with a convention that will make sense later. For example, you might have Flat1\_001.fit, Flat1\_002.fit, Flat2\_001.fit, Flat2\_002.fit, Flat3\_001.fit, Flat3\_002.fit, etc.  While here, I like to confirm where saturation is on the camera. Take the paper off if you like and dial in a much longer exposure. Mouse-around the image and see if you can read off values of 65535 (the maximum possible in a 16-bit camera). If not, increase the exposure to something like 10 s. If you still can't get to 65535, note the approximate maximum you can get to. This will be useful for estimating the approximate full-well (technically the maximum number of electrons you can record).  A few things of notes on the flats. First, you don't have to worry about dust motes too much. We can work around them by either ignoring them or by cropping around them. Second, some sensors may behave oddly with no lens attached or with a very low f-ratio lens attached. Most don't, but some do want a reasonable light cone. Here, using an SLR lens or your telescope will be needed. If your flats look reasonable, don't worry about this though as most sensors are fine. Third, make sure you're capturing your data in a raw format. If you've got a colour camera here, we don't want it to be a de-Bayered color image here. *Analyzing your Images: Basic Specifications* Now comes the fun part - seeing just how your camera's behaving. We'll cover a range of measurements, starting with the one that's most annoying. We do this not only to get it out of the way, but also because it's what gives us the ability to convert from simple intensity units (ADU) into actual electrons. I'll be using test data collected for a [review](http://www.astrophotoinsight.com/) of the Atik 314L+ I'm working on right now as an example. System Gain The system gain of your camera is the conversion rate between the raw numbers you get out of the camera (ADU or Analog Digital Units) and actual electrons. Knowing it helps you interpret the other measures as you get to express things like read noise in real units (e-) rather than in arbitrary units (ADU). It also gives you an assessment of just how many electrons you can record (which is an estimate of the full-well capacity, or at least places a lower-bound on the full-well capacity of the sensor). There are two ways to calculate the system gain: a quick and dirty one and a more involved one. I favour the more involved one described by [Tim Abbot](http://www.ctio.noao.edu/~tmca/CCD/docs/cookbook/top.html) as it's more tolerant of errors (a very similar one can be found on the [Apogee CCD University](http://www.ccd.com/ccd105.html) page).  If you decide you want to do the quick and dirty one, you only need a pair of flats and your master bias. The formula you need to compute is:  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m7bf41bbb.png  where **var** is the variance, here of the difference image between your two flats, and **mean** is the mean of the image (here of the sum of the two flats). Since the average signal in *Flat1* is really the average signal in *Flat2*, you can simplify this into:  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m593e4fb4.png  You can compute this with ImageJ, but we're going to take the longer route here. We're going to do this because any issue you may have with either of your flats will drastically throw off your estimate of the system gain without giving you any way of knowing there was an issue.  The longer route is really just an extension of this shorter route. The shorter one is using two points to estimate a line and the longer one is using several (based on the number of pairs of flats you took). It's really not so bad to do the longer route:  First, start a spreadsheet with two columns. Label them ***v*** and ***m***for ***variance***and ***mean***. For each pair of flats, you'll calculate a value for ***v*** and ***m***. We'll do ***m*** first.  Second, for each flat pair, calculate the mean intensity level (or median intensity level) across the whole image for one of the flats and multiply this by 2. This is your ***m***. Your image capture / processing software may give you this. If it doesn't, it's trivial to calculate in ImageJ. Pull down *Analyze*, *Measure* and a dialog will appear that includes the mean signal level in your image.  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m28a9c8bb.png  So, in my first pair of images, looking at Flat1, I have a mean of 6244. In the first entry in my ***m*** column, I'd then enter 12488.  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m39804df2.pngNext, for each flat pair, make a difference image. Start off by load both images in ImageJ. Before we actually subtract one image from another, we will add a constant value into one of the images. This is so that we can cleanly subtract *Flat2*from *Flat1* without "clipping" the data. If a given pixel in *Flat2* is 100 and in *Flat1* is 110, life is good and we have a difference of 10. If the pixel in *Flat1* is 90, however, we have -10 for the difference. These images don't allow negative numbers, though, so it will get clipped to 0. This will throw off our estimate of ***v***.  The solution to this is simple. Select *Flat1* (which may be actually called *Flat1\_001.fit* or something) and pull down *Image*, *Math*, *Add* and type in a number like 5000. (The actual value here won't matter. It needs to be big enough to cover the maximum difference between the images, though). Next, we'll subtract *Flat2* from this new *Flat1.*Pull down *Process*, *Image Calculator...* In the dialog that pops up, have one flat be *Image1* and the other flat be *Image2*. Select *Subtract* in the *Operation* section.  As before, we now want to measure this resulting image. So, pull down *Analyze*, *Measure* and that dialog will again pop up. Here, we're interested in the standard deviation measure. (If, for some reason, you don't see a standard deviation value, pull down *Analyze, Set Measurements*and check *Standard Deviation*). The standard deviation is just the square root of the variance (i.e., the variance is the standard deviation squared). So, we can calculate ***v*** as just:  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m67baf723.png  When I ran this on my first pair of flats, I see that the mean of this difference image (*Result of Flat1\_001*) is 5000.43 with a min of 3948 and a max of 6052. This is good as it shows that my difference image doesn't have any zeros in in (min > 0) and it isn't clipped on the top end either. The *StdDev* column shows 212.495 here, so for the ***v*** column in my first pair of images, I'd enter 45154.  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m6466f52e.pngRepeat this process for each of your pairs of flats. You should end up with a row of numbers for each pair of flats with each row having a pair of numbers. If you like, you can, of course, have your spreadsheet do a bit of the math for you by calculating ***m*** and ***v***from the means and standard deviations given in ImageJ. As you do this, keep an eye on the *Min* and *Max* values reported when you run *Measure* to make sure that you're not hitting 0 or 65535 and clipping your data.  In the end, you should have something that looks a bit like this. Here, I've entered values from four of the flat pairs from this Atik 314L+. Next, we need to perform a linear regression analysis. All this means, is that we need to fit a line to the four points we've just created. Select your data and tell your spreadsheet program to insert a chart. When asked what kind of chart to make, tell it to make an "XY Scatter". With luck, your points will all line up nicely with each other. If visually, things look like a line, proceed to the next step. If, you've got most that form a nice line but a few that are way out of line, simply delete those points from your data. Outliers typically come about from errors in your processing or image capture process or from clipping the data (e.g., hitting the saturation point of the CCD).  Next, it's time to fit that regression line. If you select your data series in the chart by clicking on one of the points in it, you'll typically have the option to add a "trend line". Different programs let you get to this in different ways, but most spreadsheets will let you do this. What you want to do is to fit a "Linear" regression and to "Show the equation" in the chart.  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_2a866386.pngThe equation will have two parts. In the example here, it says that the regression is equal to "0.27x + 292.09". That bit before the "x" is the slope of the line (you may recall the formula for a line is *y=mx+b* - this is the *m*). That slope is your system gain. It is the number of electrons per ADU. Note, typical values for this will be between 0.2 and 1.5. If you've got a number a lot higher than this, you may have flipped your ***m*** and ***v***. If so, your y-axis will have smaller numbers than your x-axis and your system gain is 1/*YourValue*.  From this slope, you can estimate the full-well capacity (or the maximum number of electrons that can be recorded before the ADC saturates, whichever is less). Multiply your slope by the maximum intensity you can get out of your image (probably 65535, but on some cameras it'll be a bit less). Here, I get about 17,700 e-.  ***Special Note 1***: This works very well if your flats are fairly flat. If they're not and if you're vignetting a lot or if you've got a whole dust-bunny warren in the image, you may want to crop a section of the image out of each flat. If you do, make sure you are cropping the exact same portion of the image out of each flat. You can either do this by carefully watching the cursor position as you crop each image or by using a cropping tool that lets you specify where to place the crop. ImageJ's *Adjust Canvas Size* will let you do this.  ***Special Note 2:***In addition, if you've got a one-shot-color camera, the "Bayer Matrix" or "Color Filter Array" on your camera may cause issues. The problem is that each color channel can have a decidedly different mean in your flats. For these cameras, I use a tool to extract one of the color channels from the raw, Bayer-encoded image. A number of programs will let you do this (e.g., Iris, Nebulosity, Maxim DL, etc.)  ***Special Note 3***: If your spreadsheet does not have the ability to give you the equation for the line on the plot there, fear not. You can use the LINEST function, passing in the ***v*** values for the x-data and the ***m*** values for the y-data. The slope parameter returned is the number you're looking for. Read Noise The system gain was by far the worst one to do, but we've gotten it out of the way and it will now let us have the other measurements be in real numbers. Next up is the camera's read noise. Recall that every time you read an image, you have some noise. This is why even with no light hitting the sensor and no dark current (bias frames), images look different.  You can typically get a good estimate of the read noise by just taking the standard deviation of a single bias frame. So, if you open up a bias frame in ImageJ and with a bias frame pull down *Analyze*, *Measure* you'll end up pretty close to the real value. But, if you want to do it right, you need to do a few extra steps.  First, if you've not made a "master bias" from all those bias frames, make one now. Use your image processing software to stack all of your bias frames (no alignment, of course) and average them all together.  Next, load up that master bias image and three or four individual bias images in ImageJ. As in the system gain measurement, add something like 5000 to your master-bias image. Then, subtract an individual bias image from the master bias image using the *Image Calculator*. Do a *Measure* on this and look at the standard deviation. This is one estimate of your read noise in ADU.  Repeat this for each of the individual bias images. It's a good idea to either keep these images open or to save them as you'll need these (and the master bias image) later on. On the Atik 314L+ here, an individual bias frame had a standard deviation of 13.93. The standard deviation of this difference image is 13.8. As you can see, we're pretty close with the two methods. The next two bias frames I tested, when subtracted from that master bias, read 13.8 as well. So, I know this is a nice, reliable measure. Average your numbers and this is your read noise in ADU. Multiply that number by your system gain (0.27 here) and you have your read noise in e-/ADU. Here, the Atik turns in an exceptional 3.7 e- of read noise. Dark Current On many cameras, dark current can be measured very easily. If you've got a cooled camera, all that is needed is to measure the mean of a bias frame and subtract this from the mean of a long dark frame. In the Atik 314L+ I have on the bench here, the mean of a bias frame is 232.5 and the mean of a 10-minute dark frame is 234.2. That means that in a 10 minutes of exposure, my average intensity went up by 1.7 ADU or 0.46 electrons. Typically, this is specified as electrons per second, so we divide this by the number of seconds in this interval (600 seconds) and get 0.00076 e-/second. This is a very low number (and is why I've often said that regulated cooling and the use of dark frames is really unnecessary on these Sony sensors - a cooled dark frame is almost exactly the same as a bias frame).  If your camera isn't cooled or if you think there might be something odd going on (or if you just want a bit cleaner estimate of the dark current), you can do the same thing you did in coming up with the system gain. In a spreadsheet, make one column for the exposure time and another column for the mean value of the dark frame at that time. Plot time on the x-axis and the dark current value on the y-axis and again do a linear fit. The data should fall on a line. If they don't something is odd as doubling the exposure duration should double the number of photons from dark current being recorded. Note, when done this way, the Atik turns in an even lower dark current of 0.0005 e-/second. The current is so low, it's really tough to estimate! Dark Stability When you collected your images, I had you collect at least 30 1-minute dark frames. This was so that you could evaluate how much the dark current changes over time. Load up each image in ImageJ and calculate the mean (average) signal, again with the *Analyze*, *Measure* tool. In your spreadsheet program, make one column (time) and enter the numbers 1-30 in there (or whatever numbers correspond to the number of darks you took here) and enter in the mean signal for the corresponding dark frame.  Again, do an X-Y plot of these data (if you like, you can select just the mean dark value and do a simple line or column plot as the x-axis is evenly spaced). You'll probably find that the camera's dark current changes a bit early on. For cooled cameras, you'll see it drop down to the set-point or to the deepest cooling point it can muster and stay relatively stable. How long does it take to get there? This will let you know how long you should let the camera stabilize before imaging. For uncooled cameras, does it reach a relatively stable point and rise no more after some amount of use? Again, this will tell you how long you should run the camera before you expect the dark current to be repeatable. *Analyzing Bias Frames and Read Noise* At this point, you've gone through and come up with some key benchmarks on your camera. You know its system gain, its read noise, its average dark current, and how stable the dark current is. Hopefully, you've also learned some tools and are now a bit more comfortable analyzing the performance of your camera. We're now going to look a bit deeper into the camera's performance by investigating the bias frames and the character of the read noise.  Before turning to your camera, it's probably worth seeing how an ideal camera would behave, as much of what we'll be looking at here isn't as clear as a simple number. In ImageJ, we can create an ideal bias frame from a camera with a clean sensor and nothing but pure, Gaussian noise. Pull down *File, New* and enter an image size of 256x256 with a background set to black. Next, add an offset to this by entering *Process, Math, Add* and entering a value of 100. You should now have a small gray image. If you were to run the *Measure* tool on this, you'd end up with a Min, Max, and Mean of 100.  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_10497b82.pngNext, add some random, Gaussian noise to the image by pulling down *Process, Noise, Add Specified Noise*, and give it a standard deviation there of 10. Running the *Measure* tool now should give you a Mean of about 100 still, but the Min and Max will now be different - perhaps about 50 and 150 respectively. The standard deviation should be about 10 (since we made an image with a mean of 100 and added noise with a standard deviation of 10...). It's probably worth saving the simulated image at this point. Histogram of Simulated Bias Pull down *Analyze, Histogram* at this point and you should see a nice, smooth histogram of your image. Again, it will show you the mean, standard deviation, minimum, and maximum. Hit the button marked *Log* to look at a logarithmic-based histogram. All this is doing is making the y-axis (height) of the histogram use a logarithmic rather than a linear scale. In log scales, the y-axis is distorted. For example, the distance between values of 1 and 10 would be the same as the distance between 10 and 100 or 100 and 1000 (this would be a log10 scale).  The figure here shows what you should see. Keep this figure on hand as it shows what a clean image really looks like. Deviations from this are not desired. We want something symmetric and that roughly resembles the nose cone of a rocket. Of course, it can have various widths, but it should have this basic shape. FFT of Simulated Bias https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m1010e077.jpgA bright bloke named Fourier came up with the idea that any signal - be it an sound, an image, a 3D shape, etc. - can be broken down into a series of sine waves. If you were to take sine waves of all the possible frequencies and combine them, adding varying amounts of each frequency, you could build up anything. If you've ever looked at the dancing lights of a spectrum analyzer on a stereo system's graphic equalizer, what you're looking at is the amount of energy in each of several audio frequency bands. This information is being derived by a Fast Fourier Transform or FFT. What we're about to do here is to analyze not the audio frequencies in sound, but the spatial frequencies in an image. If, as you move across an image you slowly ramp from dark to bright to dark, there is some energy at a low frequency. If, as you move across, you go very rapidly from dark to bright to dark again in only a few pixels, there is energy at a high spatial frequency. Our goal here is to determine how much energy there is at all possible frequencies in the image. (Note, we never have all frequencies in an image as there is a limit on the highest possible frequency that can be in an image. The *Nyquist Theorem* tells you how high a frequency can be encoded in an image. Spatially, this is two-pixels wide.)  If you've still got your image before you added the noise around (re-make it if you don't), pull down *Process, FFT, FFT*. You'll see a black square with one bright pixel in the middle. The middle of the FFT refers to 0 Hz, or "DC", or the constant offset in the image. What this is telling us is that we could recreate your frame here by adding only a single constant to the image. It's right, as the image at this point is a perfectly even gray.  Now, run the FFT on that bias image you faked. You may need to zoom in, but what you should see should look roughly like this. A bright dot in the middle with some random noise around this. What this is saying is that you can re-create your image by adding in a constant offset (the bright dot in the middle), and a number of entirely random values of random frequencies. No frequency (other than 0 Hz) is over-represented in the image.  This is really as good an image as we can ever hope for. There will always be noise in our images, but what we hope is that the noise is entirely random. Random noise will go away when stacking frames. Noise that isn't random will not go away and will build up. Remember, that's exactly what we're trying to do with our signal. Our signal consists of spatial frequencies that we want in our image. Stacking lets these remain while the noise goes away.  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m27472ae.jpg***For Fun***: If you want to get a better handle on FFTs, try doing this. First, open an image of a normal daytime shot. It may be useful to rescale it down to something a bit smaller than full-size. Here, I've taken a shot of one of my sons, Miles, at the beach. In *Process, FFT, FFT Options*, turn on *Complex Fourier Transform*. Now, do an FFT of the image. Two FFT windows will appear, looking something like the next two images here. This is the full FFT of the image. Select one of these and pull down *Process, FFT, Inverse FFT*. You'll now end up with an exact replica of the original image (top row, right). You took your image, converted it into the Fourier domain (into a frequency and phase pair of images) and then took that Fourier representation and converted it back to an image. Pretty slick eh?  In the next row here, I blanked out portions of the frequency image. By doing so, I'm cutting out a range of frequencies in the image. Pixels closer to the middle of the frequency image are lower and those closer to the edges are higher. So, here, I've cut out the higher frequency components. In one, I cut out a lot more than the other. The inverse FFTs of these restricted-frequency images now look a bit softer don't they? That's the loss of the high frequency detail. See how much you can remove before the image starts to degrade. Think this could be a good way of compressing, smoothing, or sharpening your images? https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m3685298e.jpgAnalyzing your Cameras Read Noise Frame If you've not looked at FFTs before, I don't expect this quick introduction will have you feeling like you've mastered the ideas. Hopefully, at this point you have some ideas what to look for. There are a number of good descriptions of this on the web with the one at[QSI](http://www.qsimaging.com/ccd_noise_interpret_ffts.html) being a particularly good example for us. A perfect FFT will show a bright dot in the center and simple noise elsewhere. If there are bright dots or lines elsewhere in the image, it means there are spatial frequencies in the image. That is, there is structured noise. Our goal here is to examine this noise and to determine just how repeatable the noise is. If it's repeatable, it's removable with things like bias frames and/or dark frames.  Go back to (or open up) the master bias frame you saved before and one of the images you made by subtracting an individual bias from that master bias back when we were measuring the read noise. If you run an FFT on the master bias frame, you certainly may see something that doesn't look ideal. Here, for example, is the master bias frame from the Atik and its FFT.  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m57d1db3c.jpg  You can see in the average bias that there is an odd banding on the left side of the image. The bias stack here is stretched incredibly as the total swing in the image from the dark bands to the light is about 4 ADU (on a full 16-bit scale). Likewise, the histogram shows that it's not the perfect shape (the fact that the histogram is made up of just a few spikes, though, shows that the variance in this bias frame is extremely small). Nevertheless, something is here. Something happens during the readout of the sensor to cause this slight variation in the intensity level. Since we're seeing this in a stack of 200 bias frames, odds are this is something that exists in the same place in each bias frame (or we'd never have seen it build up). If it is there in every frame, it'll come out of our light frames by subtraction. If not, or if there is anything else that is in the bias that varies from frame to frame, we'll see this in our read noise frame.  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m6e088223.jpg  The image you calculated before - this master bias minus a single bias - is a read noise frame. What is left over in this subtraction is what the camera is doing differently each time it reads the image. What it's doing the same each time got subtracted away. *This is what it does differently each time and what will show us the "fingerprint" of the camera as it were.*  So, instead of running on the master bias frame, have a look at the histogram and FFT of this difference image - your read noise image. On the Atik 314L+ here, visually, the read noise image looks very clean. Those bands have disappeared and we're left with something that looks like pure noise.  https://www.cloudynights.com/images/snr3/CCD_SNR3_html_m1061516c.jpg  Looking at the read noise histogram, we see excellent performance. There are no clear "shoulders" to the histogram and overall it has a good shape. It's not perfect, as if you squint there is a hint of a "tail" on the right, but this is excellent. We can see just how good it is by again creating a blank image, adding an offset, and adding Gaussian noise to match the values in the camera's image. I've included one here as a sample (note, if you do this, make sure the range from Min-Max is about the same in your simulation as it is in your camera's image or the histogram will differ considerably in width). Having tested a lot of cameras, I can say without reservation that this one is very, very good and there is nothing to complain about here.  Next, we can turn to the FFT of the read noise frame. Here, on the left, we see the read noise frame itself and on the right we have its FFT. As noted, the read noise frame looks very nice and smooth and it's clear from the FFT that there is nothing periodic about the noise. There are no bright lines, extra dots, etc. in the FFT. If one zooms in on it, the central dot is clearly visible (as it must be), but there is little else in the image.  Thus, we can conclude that this camera's read noise performance is excellent. The histogram is excellent and we'd be reaching here to find anything wrong with the camera. The FFTs show that the read noise is nicely random and there are no large patterns that will easily detract from the image. *Conclusions* This certainly was a long entry here and I hope that at least after several attempts, you've made it here to the end. While long, we covered a lot of ground. We covered how to get critical basic performance specifications on your camera that you might have thought were well beyond your reach, yet only required very simple tools and math. We also covered how to go deep into the analysis of your camera's electronics to see what might lie deep in the noise, but that might build up to hurt your final image.  We're not quite done with SNR here yet. We still have topics to cover like how what we know about SNR now should influence things like how we choose an image scale and what implications this has for the infamous f-ratio "myth". At this point, what I'd like to do though, is to hear from you, the reader. What parts of this haven't made sense? What questions do you have on this? I'm sure you've got questions, so drop me a line either here on the forums or by direct e-mail. I'll try my best to answer them and to shed some light on things in an upcoming entry. |

# Image Sampling

Oct 05 2009

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| We're returning to the issue of signal-to-noise in this installment a bit, but it'll take us a chunk of the column to get there directly. The topic in this round is image sampling and what I want to impress upon readers is that there is a tradeoff here. This tradeoff is between photons per pixel (our signal) and spatial resolution. We play this tradeoff all the time, but many don't really think of it as a tradeoff. Whenever we choose a pixel size for our camera, a focal length for our scope or f-ratio, a binning factor, the use of a reducer, etc., we're taking a position in this tradeoff. The goal here in this installment is to understand the trade-off. For the short answer, I'm going to argue that there's little use in going below ~1" / pixel (at least for mere mortals) and that doing so can come at the cost of a bit of SNR (of course it also often comes at the cost of a reduced field of view). So, I'm going to suggest that if you're running a 12 megapixel DSLR with 5 micron pixels on a 12" f/10 scope with its 3 m focal length, you may to want to reconsider things (as you're running at about 0.34" / pixel - a resolution your skies probably don't support and that may be costing you). SNR Recap (on the one hand) Recall from [Part 2](http://www.cloudynights.com/item.php?item_id=1973), of the series that the Signal part of SNR is:  *TotalSignal = Duration \* (Target + Skyglow + Dark)*  and that the Noise part of SNR is:  *Noise = sqrt(Total\_Signal + Read\_Noise2)*  We can put this together and talk about the SNR of our target in any given pixel as being:  https://www.cloudynights.com/images/craig-oct/CCD%20SNR4_html_13828ce6.png  Now, our *TargetSignal* here is the number of photons we capture from our DSO and the *TotalSignal* is the total number of photons captured (those from the DSO plus those from the sky and those from the dark current). This is the SNR for the target in a given pixel. When considering SNR, there comes a point of diminishing returns and for bright objects like foreground stars, the SNR is always high enough that we needn't worry about it. For cores of galaxies and such, a similar case can be made. We have plenty SNR and a 50% boost in the SNR isn't going to be noticed.  Where we do notice a real boost in the SNR is, of course, on the dim end of the scale where the SNR is a lot lower to begin with. Here, our*TargetSignal* is low and its value is often getting close to the noise. This is where we need to be most concerned about SNR.  So, when thinking about a tradeoff between spatial resolution and SNR, we should keep in mind:   1. Signal comes from photons from our target. More photons, more signal, better SNR. When we're talking about SNR at the level of a pixel in our image, we're talking about photons hitting this pixel and this pixel only. 2. One source of noise is shot noise and with brighter skies and longer exposures, this will come to dominate the noise term. 3. The other source of noise is read noise and with darker skies (or very low overall photon flux rates as you get in line-filter imaging) and shorter exposures, this can be a substantial part of the noise.  Spatial Resolution: A Visual Analogy (on the other hand) On most nights, your scope may do well looking at the moon or Jupiter at 100x (if it doesn't, consider a new scope). On a good number of nights, you may be able to push it up to 250x and be getting a good bit more detail than you would see at 100x. But, you'll be noticing about now that the image is dimmer. Why? Well, the same number of photons came off of Jupiter and went through your scope's objective at both powers, but you've now spread the light out over a bigger area. The same number of photons spread over a larger area means fewer photons per bit of area (or per square arc-second of sky).  Now, if you've got, say, an 8" scope, in theory here you should be able to push it up to 400x using the classic "50x per inch of aperture" rule. A 12" should get you go 600x, etc. How many nights can you actually do this and see more than you could at 250x? My guess is the answer is not many. Sure, on those perfect nights you can, but most nights, the atmosphere is blurring the image enough (the "seeing") that there just isn't enough resolution in the data - in that image coming from Jupiter - to get you anything extra.  *The atmosphere is providing a spatial filter, blurring the image. It limits the resolution you can ever get. Your scope provides another spatial filter, also blurring the image. Both cut out the high frequency details and limit what can be resolved (aka, limit your spatial resolution). For your scope, the aperture provides a limit based on diffraction that you'll never exceed. Even if the spatial frequency you want to capture is below this limit, the atmosphere itself can blur things enough to limit your ability to get this.*  For anyone who has tried to stick an eyeball to a scope on something bright, the above should be obvious by now. But, let's now extend this just a bit. Your eye will integrate information for about 0.1 second. Think of leaving the camera's shutter open for 100 ms as a decent analogy. Now, the seeing that causes very rapid, small scale distortions hits your eye's resolution, but seeing that happens at a slower scale (that is often larger, causing shifts of the image more than fine scale distortions) doesn't really hurt your eye's resolution. But, if you imagine keeping your eye still for 5 minutes and exposing (integrating information) over that entire time, you can bet this would limit the resolution of your eye.  Well, that's what's happening when you image. You've got the camera fixed in the same point in space (hopefully) and you've got that shutter open for a long time. Without any form of adaptive optics, there is no hope for taking the seeing factor out of your image. If you've got skies like most of us, on a pretty good night, you're looking at 2 to 3 arcseconds of blur that will be imposed on your image. That means, at best - if your guiding is perfect, if your focus is perfect, if your collimation is perfect, and if your skies are pretty darn good - you've got a 2-3" blur imposed on your image. Sure, some folks will have skies that are a bit better than this, and some nights you will too, but by and large, we've got to consider the fact that there's a good 2" or more blur that the atmosphere is giving you. Sampling and Image Scale Defined When the image is focused on our CCDs, it's a nice, smooth analog image. Under perfect conditions, you'd see nice Airy disks around the stars (at least for a moment when the skies are stable). Of course, we'll never see those Airy disks in our long-exposure images (that darn atmosphere again), but let's think for the moment about perfection. What would hit our CCD would look something like the image shown here on the left. This is the output of Aberrator showing a double-star (sampled at its default of 0.1"/pixel, 8" f/5, perfect scope, 3" separation).  https://www.cloudynights.com/images/craig-oct/CCD%20SNR4_html_m3aeb8285.jpg   |  | | --- | | *Aside: The only reason the two on the right look blocky here is that I've blown them up to match the original image's image scale. If you zoom way out so that those blocks are pixels on your screen, they look like nice stars. For very faint stars, even if you're running your scope at a very high resolution, you'll still see things looking blocky as the edges of the stars get lost in the noise and only the peak shows through. If you were to look at the audio waveform coming off of your CD player, you'd see a similar, blocky representation of the waveform. Instead of smooth waves, you'd see discrete steps. All digital signals will have this as we're quantizing (turning into discrete numbers, aka digitizing) the data. What happens in your CD player though is that there is a "lowpass filter" that cuts out the very high frequency information (higher than you can hear and higher than the sampling rate). This turns those discrete steps or blocks into a smooth waveform. Why? To actually make hard edges like stair-steps takes large amounts of very high frequency information. Remove that and you're left with the lower frequency information only (still at the limits of what we can hear), which is "smoother". The more you smooth a waveform or an image, the more high-frequency bits you're removing. By resampling this image with a bicubic filter, I'm saying that there must be a smoothness to it. That's why it can reconstruct things so well (as, in truth, the real image has a smoothness to it.) If you can have this analogy in mind and think of it in terms of spatial resolution in your image, the concepts covered here may make some more sense.* |   In the middle and the righthand panels, I've resampled this smooth star image into something our CCDs might record. Since we have individual pixels on the CCD, the recorded image will look a bit blocky. Just how blocky it is depends on the *sampling rate*. One's first reaction will be to say that we want the one on the left or perhaps the one in the middle. That is, we want our stars sampled very well so that they don't look like blocky squares. To a real extent, this is true, but there is going to be a tradeoff here and before we jump to running at as high a resolution as possible, we need to consider what we're gaining and what we're losing. By way of a preview in where I'll be going with this, consider the inset images above. I took those same blocky stars and just resampled them up to the original resolution (bicubic resampling in Photoshop).  The image scale of your rig is determined by two factors: 1) the focal length of your telescope, and 2) the size of pixels on your camera. We can use these to compute the image scale in arcseconds per pixel (assuming your focal length is given in millimeters and the pixel size in microns) as:  *ImageScale = 206.265 \* PixelSize / FocalLength*  For example, my Borg 101 ED scope when run at f/4 has a focal length of about 400 mm and my QSI 540 camera has a pixel size of 7.4u. This leads to an image scale of 3.8" / pixel. So, every pixel is covering 3.8 arcseconds of sky. Were I to run this same camera on the Celestron C8 I have here (at prime focus), I'd be at 0.76" / pixel. If you don't know the image scale for your various rigs, put this down and go compute it now.  Once we know this, we can easily compute the field of view (FOV). It's just the sampling rate times the number of pixels in each direction on our sensor. My QSI 540 has a square chip with each side having 2048 pixels. So on the Borg there, I'm at about 7782 arcseconds or just over 2.1 degrees of sky in each direction. Seeing Limits Places a Lower Bound on Useful Image Sampling Scan across various Internet sites and groups and you'll see a number of discussions on what is the "optimal" image sampling. To begin with, there is no one optimal value. If your skies might permit a sampling at one rate but covering the target requires a lower rate, well the lower rate is more optimal than the higher one. However, if the target is small, running at that higher rate will be better. But, the point of most discussions on this is that there is a limit to how well we should sample the image. That is, you won't gain anything by going lower than a certain number of arcseconds per pixel when sampling your image.  One nice treatment of this is [Stan Moore's page on pixel sampling](http://www.stanmooreastro.com/pixel_size.htm). In it, he describes things in terms of pixels per FWHM and he suggests that there is a resolution loss if you're at 1.5 pixels per FWHM and that running just a bit over 3 pixels / FWHM (3.5 seems to be a value he likes) represents about all you're going to get. So, if we plug in skies with a FWHM of 3", this leads to a value of just under 1" / pixel. Going beyond that means you're not just sampling the image well, you're really oversampling the image. That is, you're not gaining anything in terms of resolution by going at a higher rate (at the end, he gives a range of 0.5 - 1.5" / pixel for this limit which will depend on your seeing, your tracking, your gear, etc.).  Overall, I'm in agreement with Stan Moore. While he tends to put this forth as "if you want the most resolution, go for ~1" / pixel and don't skimp out at say 2" / pixel", I tend to think of it in the reverse. That is to say that there's no reason for most of us to sample at rates much higher than 1" / pixel (aka with image scales much lower than 1"/pixel) and that really, much of the time even this isn't going to buy you a heck of a lot of actual resolution. One of us is a glass-half-full and the other a glass-half-empty approach (not sure which is which, but I think I'm the empty one).   |  | | --- | | *Aside: Another way to think about this, if you like, is to imagine (or actually print this out and do it) looking at a test chart that has lines at progressively finer and finer gradations. Norman Koren has some*[*excellent ones*](http://www.normankoren.com/Tutorials/MTF5.html#newchart)*you can print and use as targets. Now, if you image this, you'll find that there is a point at which you cannot resolve line pairs anymore. This is your spatial resolution limit. If you were to place the target across a grassy field, you'd be able to resolve a finer difference than if you place the target across a road or parking lot. Why? The blur you get from the heat rising off the road will limit your ability to resolve the lines. Try it with an eyepiece if you like to see what you're up against.* |   You'll find this advice in numerous places. One bit of good coverage is in [Apogee's CCD University.](http://www.ccd.com/ccd113.html) There, they suggest as a rule of thumb dividing the typical seeing by 2 and [Starizona](http://starizona.com/acb/ccd/advtheorynyq.aspx) has the same suggestion. This isn't to say that's as best as you could possibly do, but it's getting the lion's share of the resolution your skies will afford. For 3" skies then, this would amount to 1.5" / pixel. Heck, in a review I just read by Clay Sherrod of the PlaneWave CDK-17 (Astronomy Technolgy Today, v3(4)), he said that when the scope is run at ~2000 mm of focal length (f/4.7) it "does not match well" with the 7.4u pixels on his SBIG ST-2000 camera but that it is "an incredibly good match" when binned 2x2. Unbinned, the sampling rate is 0.75 " / pixel and binned it is 1.5" / pixel. You also see professional observatories do this. The 8.2 m [Subaru telescope](http://www.subarutelescope.org/index.html) at the [Keck Observatory](http://www.keckobservatory.org/) runs at 0.2" / pixel, certainly a lot higher than 1" / second. They do have 261 robotic fingers morphing the shape of their mirror for real-time active optics to help considerably. They also have skies that, [quite often, are at 0.4" FWHM](http://www2.keck.hawaii.edu/optics/imagequality/iq.htm). Even with active optics, they're only sampling at about half a good night's FWHM of seeing. So, there's some reasoning or at least tradition behind this rule of thumb. Demonstrating Seeing and Sampling's Effects It's one thing to hear these ideas discussed and it's another to really see the effects. As mentioned in previous articles in this series, I've written a CCD simulator that tries to mimic what a CCD does in building an image. It takes an effectively perfect image (Hubble's M51 in the form of a 420Mb FITS file), adds skyglow and seeing (modeled as a Gaussian blur which, for long exposures is a reasonable approximation), pixelates the image, adds read and shot noise and quantizes it, all according to the well-known models of basic CCD behavior outlined in the first two parts of this series. By using this, we can see what the effects of various sampling rates and seeing conditions have on an image to get a feel for what we should expect and for how seeing and sampling rates interact.  Here, I've used the simulator to demonstrate the effects of 2-4" worth of seeing when sampled at 0.5" - 3" / pixel under otherwise ideal conditions (perfect camera, perfect optics, perfect tracking, and perfectly dark skies).  https://www.cloudynights.com/images/craig-oct/CCD%20SNR4_html_m791f4c70.png  We can see, of course, that overall, if you've got 2" FWHM worth of atmospheric blurring, you're a good bit better off than if you've got 4" FWHM worth or blurring. In addition, there's a solid gain in sharpness in the 2" FWHM condition as you move from 3" / pixel of sampling down to about 1.0" / pixel with perhaps a touch more to be gotten out of the 0.5" image, but only a touch. In the 3" FWHM condition, I'm not picking up any more detail below 1.5" / pixel and in the 4" FWHM condition, I'm not picking up any below 2" / pixel.   |  | | --- | | *Aside: Oddly enough, you can even argue that something like the 2" / pixel condition in the 4" FWHM looks better than the 0.5" / pixel condition. Why might this be? Even though the simulated camera has no read noise (it's a perfect camera in this regard), there is quantization error as the image is "digitized" into a 16-bit signal. The lower the signal gets, the more prominent this error, and read noise's error, become in the image. We'll return to this a bit later, but it starts to show the downside of oversampling.* |   What this is saying is that a spatial blur (here imposed by my simulated atmosphere) limits the effective maximum resolution in the image. We're not gaining anything by running that 4" FWHM image at 0.5" / pixel. We just don't have the spatial frequencies in the image, so there's no point in sampling at a rate that would record the high frequencies (that lead to our sharp edges).  You can get a feel for this in images yourself. Here, I've taken a shot (another Hubble shot), and blurred it by 2 pixels in Photoshop. This cuts out high frequencies in the image and softens the image a touch. I then copied the image, shrunk it down to 75% of its original size, and then enlarged it back to the original size.  https://www.cloudynights.com/images/craig-oct/CCD%20SNR4_html_6e90b53.jpg  If I'd done this with the image before blurring, I'd have seen a clear difference between the original and the one I shrunk and resized. By shrinking the image, I'm cutting out high frequencies (as I've now sampled the image at a lower resolution). So, when I blow it back up, those high frequencies will be gone and the image will be softer as a result. But, if I soften it ahead of time, as you can see here, there's no loss in sharpness in the shrink and re-enlarge. I never had the spatial resolution in the image to begin with, so it wasn't lost when I shrunk it. After the initial 2-pixel blur, I was now oversampled, so I could afford to do this shrink and re-enlarge without losing any detail. So, if my image were blurred like this already, there would be no need to store it at this 100% size. If I had it in this 75% size wanted it "bigger", I could just blow it up and it'd look just as sharp as if I'd had it at 100% in the first place.   |  | | --- | | *Aside, if you want to impose a spatial blur of a known FWHM in Photoshop or ImageJ, you need to keep in mind that Photoshop's Gaussian blur (and ImageJ's) specify the blur in terms of the standard deviation or sigma of the Gaussian. There's a nice formula we can use though to convert the two as FWHM = 2.35 \* sigma (or sigma = FWHM / 2.35). So, if our image is at, say 1.5" / pixel and we want a FWHM blur of 2", this blur's FWHM is 1.25 pixels (2 / 1.5). To get Photoshop or ImageJ to give a FWHM of 1.25 pixels, we'd use the Gaussian blur tool and enter in a "radius" value of 0.53 (sigma = 1.25 pixels FWHM / 2.35)* |   For those of you that want actual stars from an actual telescope on an actual camera, here is some data from a quick test shot comparing unbinned and binned (0.75" / pixel vs. 1.5" / pixel) images from a Celestron C8 on my QSI 540 (this is a small crop of a small galaxy in the frame). The raw data were stacked (no pre-processing, so forgive the hot pixel trails) and stretched linearly to match the histograms. On the left, we have the unbinned data and in the middle, we have the binned data. You can see the more pixelated stars in the binned data since the image was enlarged to equate the image size by a simple zoom. If we actually resample the middle image to the original resolution and make the two images have the same pixel count, we have the image on the right. This sure doesn't look like a two to one difference in resolution to my eyes. There may be a touch of a difference between the one on the left (native 0.75" / pixel) and the one on the right (acquired at 1.5" / pixel and resampled to 0.75" / pixel), but it's certainly not huge.  https://www.cloudynights.com/images/craig-oct/CCD%20SNR4_html_m6f860612.png  *Do not confuse this and think that I'm saying we should always run binned or that nobody will ever see a difference between binned an unbinned or between 0.75" / pixel and 1.5" / pixel. What I'm saying is that here in my skies with my gear, the spatial resolution of the image hitting my CCD on this night needed to be only sampled at ~1.5" / pixel and that going down below this to 0.75" didn't buy me much of any practical significance in terms of resolution.* If I had better skies, better optics, better focus, better hair, whiter teeth, and six-pack abs, perhaps I'd be seeing a bigger effect. Heck, on some nights I do see a somewhat reasonable boost going a bit below 1.5", but most nights its just not there and 1" / pixel would really be oversampling.   |  | | --- | | *Aside: What does binning do? Ideally (if your binning is happening inside the hardware of the camera), a group of pixels on your CCD have their charge combined before being read out. For example, in 2x2 binning, a total of 4 pixels have their charge combined. This will, of course, cut your spatial resolution in half. For it, you can get a boost in the maximum dynamic range and you get a bit of a boost in the SNR. You don't double your SNR, however. The shot noise from the target, the dark current, and the skyglow is still there and since it's driven by the signal intensity (which went up 4x), it's going up as well. Where you can get a real win is in the fact that in this bigger pixel, there is only one element of read noise. Unbinned, each pixel will have independent read noise added. When the signal is very faint and when your noise is dominated by read noise (you've got dark skies or are running with a line filter), binning can help boost the SNR on extended targets. But, as the target brightens overall (you're getting away from the read noise and quantization error) and as the noise shifts to shot noise from things like the skyglow, binning isn't really boosting the SNR much. Sure, it looks "brighter", but stretch the unbinned image and you can brighten it up. Some of the same reasons why oversampling can hurt you are the reasons why binning can help you. But, some of the same reasons that don't make oversampling too horrible, make binning a bit less useful than many might think.* |  The Downsides of Oversampling So far, we've been coming at this by saying that there is a limit to what you can expect out of your system and that there's little point in going below this and sampling at a higher rate. At the outset, though, I said that there is a tradeoff, however, which means the higher sampling rate is coming at a cost. Just what is that cost?  There are three costs to consider here. One should be rather obvious and that is that if you're sampling at a higher rate by changing the focal length of your scope, you're going to cover less sky. I've got 2048 pixels and run at 1.5" / pixel, I'm covering 51". If I run at 0.5" / pixel, I'm covering 17" of sky. Now, there are some good number of small targets that this won't matter for, but there are a good number of larger targets for which it will matter. If you're not gaining anything resolution-wise by running at the higher rate, why sacrifice coverage of more sky? Why not give yourself more breathing room around the target? Sure, it may look smaller on the image, but that's what the crop tool is made for! (Of course, if you change the sampling by binning your CCD or by using a different CCD with the same physical dimensions but different pixel sizes, the FOV won't change.)  A second cost is more psychological with its physical manifestations coming out as self-induced hair loss. I started out giving the example of a DSLR run on a 12" f/10 system at prime focus. Here, you're looking at 0.34" / pixel. Let's say that you've gotten your guiding accuracy down pretty darn well and your RMS error in RA is under an arcsecond. You should be happy and if you're imaging at 1 or 1.5"/pixel your stars will look nice and round. Image at very high resolutions though and your residual guiding error will still come through. At the "lower resolution" (but still high enough for what your skies can support), you'd never know about this error, you'd like your shots and be proud of what you're doing. In short, you'd be happy and get to enjoy the hobby and take pride in the level of accuracy you've achieved. You've gotten guiding down well enough that everything the atmosphere allows you to have resolution wise, you've gotten and your stars are still round. You'd still have your hair (well, you would if you started with it). But, run at very high resolutions by oversampling and this all goes away. That ignorance being bliss thing goes out the window and it does so for no good reason. You're frustrated and spend nights trying to fine tune your guiding so that the stars come out round even at this level of magnification. You lose valuable imaging time and gain what? Round stars at an image scale your skies can't support anyway. For me, life's too short to worry about that. Give me the happy-imager recipe.  The third cost is one of SNR. If we keep the aperture of our scope constant and only change the focal length (i.e., change the f-ratio by reducing or extending it), we don't change the total number of photons going into our scope. The DSO is streaming photons from space and our scope is catching them with a bucket the size of our aperture. Running at a higher sampling rate means spreading the light from the DSO across more pixels.  Thus, each pixel is getting less light and so the signal hitting that pixel is less. Some aspects of the noise (e.g., read noise) will be constant (not scale with the intensity of the signal the way shot noise does). Thus as the signal gets very faint, it gets closer and closer to the read noise. As we get closer and closer to the noise floor, the image looks crummier and crummier. Doubling the focal length (aka one f-stop or doubling the sampling rate) will have 25% as much light hitting the CCD well, meaning we will be that much closer to the read noise. If the exposure length is long enough such that the bits of the galaxy or nebula are still well above this noise, it matters little if at all. But, if we are pushing this and just barely above the noise (or if our camera has a good bit of noise), this will more rapidly come into play. (Furthermore, who among us doesn't routinely have other faint bits that it'd be great to pull out from the image?)  Please note, that none of what I am saying here contradicts Stan Moore's "[f-ratio myth](http://www.stanmooreastro.com/f_ratio_myth.htm)" page. He makes this same point and if you look closely at the images on his site, the lower f-ratio shot does appear to have less noise. As noted, itâ€™s not "10x better" (which some people who say f-ratio is all that matters would argue), but itâ€™s not the same either. Stan argues that, "There is an actual relationship between S/N and f-ratio, but it is not the simple characterization of the â€˜f-ratio mythâ€™." What I'm arguing here is to try to make clear that other side. F-ratio (and therefore image sampling) doesn't rule the day and account for everything, but it also isn't entirely irrelevant.  Here I’ve taken some data from [Mark Keitelâ’s](http://aajonahfish.com/f-ratiomyth.htm) site. Mark was kind enough to post FITS files of M1 taken through an FRC 300 at f/7.8 and f/5.9 and to give me permission to use them. I ran a DDP on the data and used Photoshop to match black and white points and to crop the two frames.  https://www.cloudynights.com/images/craig-oct/CCD%20SNR4_html_m9acd2aa.jpg  To get a better view, here is a crop around the red and yellow circled areas. In each of these, the left image is the one at f/7.8 and the right at f/5.9 (as you might guess from the difference in scale. Now, look carefully at the circled areas. You can see there is more detail recorded at the lower f-ratio (lower sampling rate). We can see the noise here in the image and that these bits are closer to the noise floor.  https://www.cloudynights.com/images/craig-oct/CCD%20SNR4_html_206909e9.jpg  Again, the point is that it’s incorrect to say that the f-ratio rules all and that a 1” scope at f/5 is equal to a 100” scope at f/5, but it’s also wrong to say that under real-world conditions, it’s entirely irrelevant. For a given aperture, f-ratio and image sampling rate are synonymous.  Is it a huge effect? No, but it's one that will be present to varying degrees and one that can hit you where it hurts. If you're running with a line filter and trying to get that faint H-alpha image and are already pushing to get 5, 10, or 15 minute shots to show much of anything, you're running down near the read noise. If you're down near the read noise, you're SNR in that part of the DSO is very low. Spreading the light across more pixels will drop the SNR and make that part look crummy. Run at a lower resolution (smaller f-ratio, lower sampling rate, etc.) and you're getting more photons to hit that same CCD well, getting you further away from the read noise.  Therefore, for the same exact reasons why f-ratio matters some, image sampling matters some when it comes to target SNR. As noted in the *Aside*above, binning has a very similar effect here. Under the right (or maybe that should be "wrong") circumstances, your SNR will go down as you oversample. Taken to extreme levels of oversampling (e.g., 0.1" / pixel) you darn well better be able to expose individual subframes long enough to get your signal well above this. Conclusions Hopefully, at this point, you've got a good idea not only of what image sampling is, but also that there is a bit of a tradeoff when trying to pick an image sampling rate. I'd like to leave off with a few basic conclusions:   1. Sampling rate is defined by the focal length of your telescope and the size of the pixels. Adjusting either (changing scopes, using focal reducers, changing cameras, using binning, etc.) will change the sampling rate. 2. There is no one, perfect, thou-shalt-always-use image sampling rate. 3. Even if you decide upon a target sampling rate of something like 1" / pixel, don't go making drastic changes to your system if you can currently hit 0.9" / pixel or 1.1" / pixel. You won't notice a difference in resolution or SNR. Values here are guidelines and they're not hard and fast numbers. 4. Your skies, equipment, and ability to get the most out of the equipment are going to place a bound on how much resolution you can get out of your image. When starting out, you and the equipment may establish that boundary. Once you've got focus and guiding down well, though, the skies are likely going to be the determining factor. 5. Running at sampling rates much below a half or a third of your skies' FWHM isn't going to bring in much if any more resolution in your shot. Other things blurred it enough before it even got to you that there's just nothing more you can wring out of it. For most of us then, a value of about 1" / second will be as high a sampling rate as we should use when trying to do high-resolution work. For a lot of us, for normal imaging, you won't be losing much (if anything) by even running at 1.5" / pixel. (I, personally, use 1.25" to 1.5" as a good target sampling rate for small targets and a lot more than this for wide targets. My favorite rig runs at over 3" / pixel). 6. Running at very high sampling rates has several downsides. You're FOV may be more limited, tracking errors are more visible, and SNR can be reduced. 7. For many cameras, these points taken together suggest that your scope's focal length can be limited to ~1500 mm with ~1000 mm being a fine target value for high-resolution work. Most DSLRs have pixel sizes of about 5 microns. Many dedicated CCDs have pixel sizes of 6.4 or 7.4u (a few go up to 9 u). If we take 6.4u then as a fairly typical value we find that 1" / pixel is reached at 1320 mm of focal length. 1.5" / pixel is down at 880 mm of focal length. Before you put that DSLR onto a 3000 mm scope, be aware that you're solidly over on the other side of this tradeoff, asking for resolution you almost certainly don't have. In the process, you've given up a few good things: FOV (assuming you can change focal length to affect sampling), ease of guiding (and perhaps sanity or some hair), and a bit of low-level SNR. Instead, start looking at shorter focal length setups or cameras with much larger pixel sizes (the former are much easier to find). Imaging will be a lot more fun and you won't actually have given up much if anything in the resolution department. |

# Image Sampling Myths - Part 5

Jan 26 2010

|  |  |
| --- | --- |
| In the [previous entry in the series](http://www.cloudynights.com/item.php?item_id=2042), I went over a number of aspects of image sampling. In it, I argued that running at rates much more than a half (or perhaps a third) the FWHM of your skies wasn't going to buy you anything more in terms of resolution in your shots. At the same time, I argued that you're going to get a hit in your SNR. I purposely didn't come down hard on the "f-ratio myth". I danced around a few issues there and left things pretty broad by saying that "f-ratio doesn't rule the day and account for everything, but it also isn't entirely irrelevant." A great thing about Cloudy Nights is that there are forums there and readers can call you on it. In particular, I have to thank three individuals for pushing me on this: Sergi "xatamec", Frank "freestar8n", and Mitch Cox. I'd been thinking of developing this a bit more but these three get the credit for giving me a rather large shove. Thanks gents!  So, in this entry, we'll cover three things that have seemed to become dogma in our hobby, each of which has some problems that lead me to reject the simple mantras. The first is on Nyquist and as it pertains to image sampling and the second is on f-ratios. Both expand on [Part 4](http://www.cloudynights.com/item.php?item_id=2042), so if you've not read that back, you may find it useful to do so before continuing. The third is really an extension of the second and covers how much of this debate is driven by film vs. CCD. Myth 1: The Nyquist Theorem doesn't apply to our images The first thing on the docket is the notion that the Nyquist Theorem (and its sampling limit) applies only to things like one-dimensional audio streams and does not apply to things like two-dimensional images. This is pure bunk. If you have an two-dimensional, three-dimensional, or even N-dimensional analog source of information and wish to digitize it, the Nyquist Theorem applies. It is a theorem about information. It and Fourier analyses don't give a whit about the number of dimensions they are working on.  That said, the Nyquist Theorem has a number of stipulations. It works perfectly when you adhere to these stipulations, but when we don't, it can break down a bit and we won't get a perfect reconstruction of our data. We'll come back to this a bit at the end. But to begin with, what is the Nyquist Theorem and what is Fourier analysis?  To start with the latter, in its commonly-applied form, Fourier analysis breaks a stream of information into a set of component frequencies (and complex phases). This transformation is invertible such that you can reverse it and re-create the original stream. The stream can vary along time (as we have in an audio waveform) or in space (as we move across an image in each direction). Once in the frequency (and complex phase) domain, we can look for and potentially filter out certain components. (This is similar on the surface of things to what Registax does in letting you filter or enhance different wavelets.) So, we can use it to detect things (like RF noise in your camera - see [Part 3: Measuring your Camera](http://www.cloudynights.com/item.php?item_id=2001)), to enhance frequencies (e.g., enhance relatively high frequencies to sharpen an image), to reduce frequencies, to compress an image, etc. It's really an amazingly useful concept and our modern technological world couldn't exist without it.  Fourier analysis says that we can break down any image into its component frequencies and then invert this transformation to re-create the image. I demo'ed it in [Part 3](http://www.cloudynights.com/item.php?item_id=2001) with an image of my son. The question then comes up: if we are going to digitize this, how well do we need to sample it? What is the relationship between frequency in the image and sampling rate? This is where Nyquist comes in.  The Nyquist Theorem says that in order to reproduce frequency X, you must sample the stream at a rate of at least two times X. So, for a 20 kHz tone (most readers can't hear this high anymore), you need to sample it at 40 kHz. Now, in the conversion from analog streams (e.g., music coming from a trumpet) to digital, we have to worry about aliasing. Frequencies in the stream that are higher than the sampling rate get aliased back in as lower frequencies (think of beat-frequencies if you know what they are to give a basic idea of the process. In practice, it comes through as noise.) So, we must filter them out and it's the construction of these filters that leads to CDs having sampling rates of 44.1 kHz rather than 40 kHz. Imperfections in making the filters lead to the desire to have a bit of breathing room here. Thus, in digital systems, you'll often find things running a bit higher rates than Nyquist would require. This, also, is where we have a bit of an issue with our images. Most cameras don't have a lowpass spatial filter to block the higher frequencies that will alias into the image as noise.   |  | | --- | | *Aside: You be thinking -- Hey, that Nyquist thing says that the frequencies are sine waves. What if I had a 20 kHz triangle wave or square wave? It'd reproduce that as a 20 kHz sine wave after all that Fourier stuff and conversion. Well, yes, it would. Actually, coming out of the DAC would be a 20 kHz square wave. But what makes that 20 kHz wave square are very high frequency components that turn the sine wave into a square (or near-square - perfect square waves don't exist). Remove those very, very high frequency components and you get back the 20 kHz sine wave. So, that 20 kHz non-sine wave has frequencies higher than 20 kHz. If we want to accurately reproduce them, we need our sampling rate to be a lot higher than 40 kHz.* |     **Figure 1-1**  https://www.cloudynights.com/images/craig-starks/part-5/Nyquist1.png  That concern aside, we can demo this ourselves to get a feeling for how this works using readily available software (Aberrator and ImageJ). First, I used Aberrator to make some images. The ones I'm working with here are of a double star, 2" apart stars in an 8" scope (optically perfect, upper-left of Figure 1-1). *So, this is a pretty tight double for us to be resolving in our DSO images*. Then, I brought these into ImageJ (see Part 3 for details on ImageJ) and made two copies of the image. I did a 1" FWHM blur on one and a 2" FWHM blur on the other. Recall that ImageJ uses sigma, not FWHM to specify the blur size (as does Photoshop). So, the blur sizes were 4.25 and 8.5 pixels (the image scale is 0.1" / pixel, so 2" FWHM is 20 pixels, and divided by 2.35 to convert to sigma and you get 8.5). I then did an FFT on both of them and you see the resultant loss of high frequency detail in the FFT of the blurred image. Remember, the center of the FFT image is the lowest spatial frequency and moving out from the center shows the energy at higher spatial frequencies. So, if you have 1" or 2" worth of seeing, this is what you get.  Note how you can see the image soften with blurring as we move across the top row and how the FFTs show less and less energy at higher frequencies. Note in particular how much high frequency information has been lost with only 1" FWHM of blur added to the image (the FFT is now very compacted in the center). To help us see how FFTs and frequencies work, I next took that first one (the unblurred image from the upper left) and cropped the FFT with a circular mask to cut out about half the frequency space and did an inverse FFT of it. The result is in Figure 1-2 (top row is the "input", bottom row is the "output").  **Figure 1-2**  https://www.cloudynights.com/images/craig-starks/part-5/Nyquist2.png  You can see we can crop out a lot of the frequency domain (the high frequencies) and do very little to the image. The inverse FFT of the upper-middle (shown in the lower-middle) looks very similar to the original. Those outer rings have very little energy even in this perfect image and you can't really tell upper-left from lower-middle. Cut out a lot and you can (a harsher crop of the Fourier data is shown on the right side), but the point is we certainly don't need 0.1"/pixel resolution even here on this perfect image. Not only does the FFT show the whole frequency space isn't being used (lower-left), but a good bit of that is even carrying very little information. Heck, if we cut things very hard (as on the right) we still get something reasonable.  Notice, now that the very hard one has a hard circular crop in the frequency domain about where our pure image, blurred by 1" FWHM had all of its energy (Figure X1, lower-middle). The radius of that is about 25 pixels (or about 1/10th the image). What that means is that there is no energy at a frequency higher than this. Given the energy content, 1" is all the resolution we have, so we can sample this at 0.5" and not lose anything at all. In Figure 1-3, I took that image, rescaled it in ImageJ down to 48x48 pixels (a 5.3x reduction using bilinear interpolation, putting us at 0.53"/pixel) and then blew it back up with bicubic resampling. This shows that as long as we sample the image appropriately, nothing need be lost.  **Figure 1-3**  https://www.cloudynights.com/images/craig-starks/part-5/Nyquist3.png  Figure 1-2 showed that even if we undersample our image rather dramatically (we cut out a lot of the data in the Fourier domain by that circular crop), we can still have a good amount of detail. In any case, as we get into real world situations with 1", 2", 3" or sometimes more worth of blur, running at rates higher than the Nyquist rate isn't going to buy you anything, even on highly magnified stars. Keep in mind, these two stars are only 2" apart. We're looking at the most detailed kinds of things you're looking to resolve and this simulation here is with only 1" FWHM of seeing. My skies are routinely a lot worse than this. Figure out the limiting factor in your seeing (be it skies, tracking, or driven by diffraction given the aperture of your scope), halve that and don't consider it a good thing to go beyond this. Even if you're on the other side of this and aren't "critically sampling" (or oversampling) the image, the high frequency component you've lost may not have that huge an impact on the overall image. Look again at the middle column in Figure X2. There, I cropped off some of the high frequency information before doing the inverse FFT. Can you really tell the difference between the upper-left and the lower-middle panel? After stacking a bunch of pictures would any difference you see now still remain?  So, to sum up, the Nyquist Theorem says that if we follow all the rules, we need only sample an image at a rate that is twice the highest spatial frequency we want to be able to faithfully reconstruct. We'll bend the rules a bit at times and we have things like how f-ratio affects star shapes, how CCDs are imperfect, etc. that will place other limits on our performance (ask Frank freestar8 for more details on this and be prepared to take notes!). So, if the utmost fidelity in spatial resolution is the goal, going a bit beyond the 2x rate is a good idea. For many of us, the amount of information in those very high frequencies may not be enough to justify the higher sampling rate and we do need to keep in mind that our skies are placing strong limits on what most of us can achieve. We then come back to the suggestion that if you're going for the most spatial detail possible, running at half or perhaps a third of your seeing is really all you should aim for. If you're looking for something more general-purpose (and if you don't want a hit in SNR), you should aim for something wider as we'll see in Myth 2. Myth 2: For SNR, the f-ratio doesn't matter or matters only in extreme or marginal cases. This is a far more contentious issue. So, with flame retardant suit donned, I will continue, ready to be considered disparagingly in some circles. As noted at the outset, I've not been entirely clear on it myself. This is my attempt to set the record straight so far as I see it. One thing to note at the outset is the way I see it is biased by the kinds of images I want to take. I'm not into photometry and pictures of double-stars don't get me going either. I'm a DSO guy who treats astrophotography as a technical art. I have nothing against science (I'm a scientist by day), but my goal here is to make nice images of DSOs. These are inherently extended objects, so I pay particular attention to how well we can image them. This places more weight on one side of a trade-off - the trade-off I brought up in the last installment. But, as with the last installment, the key here is that a trade-off exists.  So, if you've got the same aperture of scope, does the SNR of the image get affected by what f-ratio you're running at? Pixel-based vs. Object-based SNR To begin with, we have to come to grips yet one more time with the concept of SNR. Specifically, we need to be clear about what "kind" of SNR we're talking about. We have to do this because the popular "[CCD f-ratio Myth](http://www.stanmooreastro.com/f_ratio_myth.htm)" page by Stan Moore has a different take on what kind of SNR we should be talking about. To understand where my take splits from Stan's and to understand why I've said in the other articles that none of the text contradicts his work, we need to delve into the two definitions of SNR. In all this, we'll hit again something akin to one of us looking at the glass as half full and the other as it half empty. Since I took the half-empty one last time, I'll give that honor to him this time.  First off, we should point out where there is clear agreement:   * Stan says, "There is an actual relationship between S/N and f-ratio, but it is not the simple characterization of the f-ratio myth." *Agreed. One f-stop will not double the SNR as many would have you believe.* * He also says, "Information about an astronomical object (star, galaxy, nebula, features of galaxy or nebula, etc.) is contained in the light that falls onto Earth. That light consists of a certain number of photons per second per square meter of earth's surface. The quality of information from an object depends on how many photons are captured and measured by the instrument." *Agreed again. But, we're starting to see the point of divergence here as he is talking about "information". We should also note that the photons need to arrive at a portion of the image plane that you've got covered by your sensor.*   For Stan, the "true SNR" is "object SNR" and this "refers to the actual information content of the image". To understand this, you must think about that term "information". Think of it as how much data are really in the image or how many bits it would take to compress the image without losing anything (have a look at Shannon Information Theory if you want to read up on this). For example, a dark background with a bit of a galaxy core showing through has less information than a dark background with a nice bright galaxy showing through. If you capture more photons (e.g., with a bigger aperture) you will have more information, all else being equal. So far, so good.  Likewise, if you have the same galaxy image overall such that a thumbnail version of two images would look the same but one has a lot of detail in the arms and the other is blurred, the more detailed one has more information than the more blurry one. The more information in your image, the better. Here, again, I will agree in principle.  The problem is, by using the term "information", we've muddied the water a bit. It's held up as a Holy Grail so to speak or at least a lot more interesting than the more mundane pixel-based SNR. While it can be treated purely objectively (and we can calculate exactly how many bits of information are in an image), for us, there is often a subjective aspect to it. There is the real concept of useful information vs. useless information. What if we don't care so much about spatial resolution (spatial information) and are willing to sacrifice some of those details by undersampling a bit so that we can get a much wider field of view? We've gained a good bit of information (many more stars and galaxies in the image, for example) but lost some in the process (detail in the arms of said galaxies). Is one better than the other?  Should I really be working to maximize a raw measure of information here? I don't think so. It's a useful way to think about things, but I don't think we must diligently work to maximize this value. Heck, I'm a sucker for wide-field shots, having clearly lost some spatial resolution (gaining and losing information here in the process). I'm not sure I care where Shannon would come down on this. To me, the wider field of view's added information more than offset what I lost in terms of resolution. Were I interested in planetary nebulae, I'd sing a different tune, however.  I also don't think that it's the best way to think of SNR as it pertains to making a nice, clean image of DSOs. I agree entirely that the bigger the aperture of your scope the more photons you will capture from a given target, assuming that target fits on your CCD chip. That's a given. But, each CCD well is largely independent (and CMOS sensors are even more so). Each pixel's job is to estimate to the best of its ability how many photons hit it and it doesn't care a) whether the photons are from a DSO, the skyglow, a star, or from heat and  what the CCD well next to it is doing. It's not like all the pixels that are part of the DSO all get together to share notes and secret handshakes. Each is just a detector. This detector is described by that simple pixel SNR equation.  Now, when we make an image out of these pixels our eyes and visual systems do impose a relationship among nearby pixels. We form lines and edges and we're very good at picking them out amidst noise (read up on "Gestalt psychology" for some fun demos of this). But, as we're doing our various DDPs, curves, and levels, we're just taking each pixel's value for what it is and shifting it according to a transfer function (e.g., an S-shaped gamma, an arbitrary curve, or a power function). If the estimate of the true intensity value for each pixel is closer to the actual truth (i.e., if there is less noise), we can stretch things more before the image breaks down. That is, if the pixel SNR is higher, we can stretch it more before the image looks noisy.  Thus, it is the pixel-based SNR that I worry about. This is the SNR we have defined in the other entries in the series and it is this SNR that you will find in typical discussions of the SNR in images and in our CCD cameras. We talk about the estimate of the number of photons captured by that pixel for that pixel's area of sky (not the whole detector's area of sky or the entire DSO's area of sky) and the variance in that estimate. As someone who is interested in imaging extended objects and picking them out as cleanly as possible from the background, this is a very important kind of SNR. I care about this first and spatial resolution second. Detail in a galaxy arm cannot be had unless you can record that galaxy arm in the first place. (Note, when I have talked about "Target SNR" in other articles and here, I'm talking about the SNR in that pixel for the photons from the target. I do this to make sure that skyglow isn't part of our "signal". It's not talking about SNR for the entire target, but rather is the "Target Pixel SNR").  So, Stan and I are talking about two fundamentally different concepts as we use very different definitions of SNR. This is why I have given a wide berth to his coverage and said that none of what I have gone over has really contradicted him. We've been talking about two different things. That said, not only this has led to some real confusion, but it's not been entirely accurate. For, even when you take real "information" into account (i.e., you have critically sampled the image so you have not lost any spatial resolution - see Myth 1), f-ratio still has an effect and it's not a non-trivial one. The effect is going to come down to not only a role for the read noise in the camera, but also a role for the shot noise from the target that hasn't always been considered. The Role of f-Ratio on Pixel-SNR: The Equations Let's go back to the full pixel SNR equation (as always, having the "signal" be the photons only coming from the target and not the photons coming from the skyglow and dark current as we want a pretty picture of the target). We'll do this to drive home the point that in terms of pixel-SNR, f-ratio is going to matter. We have:  https://www.cloudynights.com/images/craig-starks/part-5/FullSNREqn.png  Ã¯Â¿Â¼ For now, let's pretend we have a perfect camera with no dark current and no read noise. We thus just have our target's signal and the skyglow's signal:  https://www.cloudynights.com/images/craig-starks/part-5/PerfectCamSNREqn.png  Now, let's consider what happens if we keep the aperture constant and we change the f-ratio by one stop. Say, we went from a 100 mm f/5.6 scope to a 100 mm f/4 scope. We've shifted from 560 mm to 400 mm of focal length. Were your SLR hooked up to these scopes, it would compensate by halving the exposure duration.  Why? Each pixel (and the whole sensor) is getting twice as much light (assuming you're shooting a flat frame or something) as it's covering twice as much sky. Run the math if you like and you'll see that each pixel is covering sky 1.414 times as wide and 1.414 times as high (sqrt(2)). You can also think of this as making each pixel physically bigger by a factor of 1.414 in each direction.  Pretend for the moment that this pixel is aimed at a relatively smooth part of some nebula. So, expanding its FOV hasn't made it now in some brighter or darker patch. It's the same basic stuff, only more of it. What happens to the SNR? Let's call this TargetPixSNR' (the ' to separate it from the first one we calculated). We get:  https://www.cloudynights.com/images/craig-starks/part-5/OneStopSNR1.png  Simplify this a touch and we get:  https://www.cloudynights.com/images/craig-starks/part-5/OneStopSNR2.png  Thus, by having a one-stop change in the f-ratio and keeping the aperture constant, we have boosted the SNR here by a factor of 1.414. We boosted the signal by a factor of 2 (which is why your SLR will halve the shutter speed), but the SNR went up by only 1.414. That said, it didn't go up by something like 1.0001. No, it went up by 41%. That's not chump change.  Note, the 1.414 here comes from it being the ratio of the f-ratios (or the ratio of the focal lengths since the aperture here is constant). It's also the square root of the boost in the amount of light. Here, we doubled the amount of light (by doubling the area of sky) and sqrt(2) = 1.414. If we'd cut the focal length in half down to 280 mm, we'd have quadrupled the amount of light and boosted the SNR by a factor of 2 since sqrt(4) = 2. Again, we can get to this by the square root of the boost in light or by just the ratio of the focal lengths or f-ratios (560 / 280 = 2 and 5.6 / 2.8 = 2). The Role of f-Ratio on Pixel-SNR: An Insanely Simple Application Some readers don't like it when things are done in "just math". Let's take an actual image but one we can control here. We want something with an actual scope and an actual camera, but we want the target to not be affected by seeing, transparency or actual contrast variations (we're looking at an even part of this nebula such that our pixel's coverage is still in the same basic intensity range despite covering more sky). One such target would be a flat frame. We can use this as a proof of concept for extended objects (again, not stars and yes, we are ignoring the obvious fact that spatial information will lost here).  Here, I took my Borg 101 ED run at its native 630 mm focal length (f/6.3) and run at its reduced focal length of 400 mm (f/4) and shot my EL flat panel using my QSI 540. The scope was focused at infinity and 20 frames were taken at 0.001, 0.01, and 0.1 seconds using each optical configuration. Images were bias corrected using a large stack of bias frames. For the f/6.3 rig, the mean intensity in the center of the image ranged from 154-13.8k ADU and for the f/4 rig, the mean intensity in the center of the image ranged from 357-31.8k ADU after this processing.  **Figure 2-1**  https://www.cloudynights.com/images/craig-starks/part-5/FlatData.png  Now, according to the math above, the boost in the SNR here should be 630 / 400 or 1.575. I calculated the SNR in my test images two ways. The quick and dirty way many would do it is to just take the mean intensity in a locally-flat area (here, a 10x10 area in the middle of the image) and divide it by the standard deviation in that area. This is a good proxy for the SNR, but it does assume the image is perfectly flat and that the sensor (post bias correction) is perfectly flat. To clean this up a bit, I did it the more exacting way as well, calculating the mean value and standard deviation for a given pixel across the 20 images I took. The former is the "SNR" and the latter the "BSNR" (better SNR) in the graph here in Figure 2-1.  On the left, you can see the SNR for the f/4 configuration is higher regardless of the image intensity or the measure of SNR. On the right, we see how much higher it is. With a perfect camera, no extra light loss associated with the reducer, no non-linearities in the detector, etc., this should come out to a 1.575x improvement for the f/4 condition over the f/6.3 condition. The mean of the SNR bars here on the right is 1.54 and the mean of the BSNR bars on the right is 1.5. I'd say the math held up pretty darn well.  So, in this semi-real-world test, a shift in the f-ratio, holding aperture constant boosted the SNR of an extended object by the predicted amount. Reducing the f-ratio by a factor of 1.575 boosted the SNR by this amount as well. *Thus, f-ratio is clearly having an effect on SNR and it's not something to be swept under the rug.* Getting Back to Information At this point, those who uphold the notion that f-ratio's effects are a myth may be rather incensed as what I've done here is to present a case with very little real information in it. A perfect image of a flat has almost no information in it. You could represent all 4 million pixels in my camera with a single value, say 1327 if the optics were perfect and there were no noise. A single value (repeated across all pixels) is very little information. ***The point of that exercise, however, was to show that when we're not talking about gaining or losing some spatial detail, the f-ratio surely does matter in our ability to estimate the intensity of that spot of the nebula.***Move over some number of pixels where the nebula is perhaps darker and we'll have a lower value. Our SNR in that pixel will be higher with a lower f-ratio. Thus, when we stretch the image to better reveal the difference between these areas, we'll have a lower-noise image and be able to show the contrast between the two areas better.  What happens when we have actual spatial information in here we're worried about losing? If we turn back to Myth 1 here, we note that oversampling an image much beyond what Nyquist would indicate isn't letting you capture any more information. You have a certain amount of information in the what is passing through the cover-glass on your CCD and that amount is dictated by the inherent blur in the image. That blur is dictated by the effects of aperture (larger apertures reduce the size of the Airy disk), seeing, tracking, etc. At this point, how it is sampled and what the f-ratio are doesn't matter.  But, we must take that information as it is passing through the cover-glass on the CCD and we must sample it and record it using real detectors. Let's pick apart the two aspects of information here: spatial information and intensity information.  For spatial information, we must consider the sampling rate. If we are at or above the Nyquist rate, we are not losing spatial information. If we are below this rate, we are losing spatial information. How much we are losing depends on how badly we are undersampling the image. Depending on how we have gotten to this undersampling point, we may or may not have gained information. If we have a sensor of the same physical size and have just increased the size of the pixels (e.g, by binning), we have lost spatial information (again, if we are below the Nyquist rate). If we have kept the same sensor and have done something like put a focal reducer in place, we have lost spatial resolution but we have gained field of view (FOV). Clearly, those extra stars, galaxies, nebulae, etc. are information. Whether it is useful information or not is in the eye of the beholder. But, if your arcseconds per pixel is more than roughly half your seeing (or whatever else is limiting your resolution), you're certainly starting to lose some spatial resolution. Whether that trade-off you've made is worth it or not (e.g., by exchanging some spatial resolution for greater FOV), again is in the eye of the beholder. This is technical art and when we consider changes in FOV, the technical bits like consideration of information-content break down.  For intensity information, as shown above, we must consider the fact that these are real-world detectors of photons and that photons behave according to a Poisson process ([see Part 1](http://www.cloudynights.com/item.php?item_id=1966)). We have read noise and we have shot noise. Ignoring read noise, if you pack N times as many photons onto a CCD well, your SNR will be sqrt(N) higher as a result of this Poisson process. How do you get more photons to pack into a CCD well? Alter the f-ratio. The aperture will determines the total number of photons from the sky that are collected. The focal length determines the size of the image and therefore how they are spread over the sensor. The f-ratio, being the ratio of these two, determines how many photons are packed into each square millimeter (or micron, or what have you).  If we go back to [Part 4](http://www.cloudynights.com/item.php?item_id=2042) and look at the equation for the image scale, we see that the number of arcseconds per pixel is based on the focal length (206.265 \* PixelSize / FocalLength). So, if we keep the focal length constant, we keep each pixel covering the same amount of sky. If we want to get more photons hitting the detector from this same amount of sky, how do we do it? We increase the aperture. Aperture wins? Not so fast, because in so doing, we dropped the f-ratio. Same focal length, but larger aperture results in a lower f-ratio. Take that increased aperture's extra photons and spread them out more now so that you have the same number of photons hitting each CCD well as you had with the smaller scope. To do that, we've had to boost the focal length, making a larger image on our CCD and having each pixel cover a smaller area of the sky. We've boosted the f-ratio back and, in fact, we've boosted it to exactly the level we were at before. The f-ratio is tracking the number of photons for an extended target perfectly. That's what it's designed to do. There's nothing magical about film or CCDs here. It's just simple geometry.  Can we claim that the only thing affecting information is aperture? No, I do not believe we can. If we keep f-ratio constant and scale the aperture we are scaling the focal length. This will allow us to gain spatial resolution up to the point at which the data coming through the cover-glass have revealed all they can and there is nothing left to be pulled from it (owing to the inherent blur). Going beyond that and we are not actually gaining information in the image (there is no more detail to observe). We are clearly losing information as well by restricting our FOV. We're also losing information by dropping the photon count in each pixel. By dropping the photon count in each pixel, we're getting closer not only to the read noise but also to the shot noise. This is what is missing in most discussions of f-ratio and SNR.  Consider Figure 2-1 above again. On the left, we see the SNR is clearly much, much better for the longer exposure. Sure, the short one is closer to the read noise, but both the two longer exposures are far away from it (and, in truth, with the 7 e- of read noise in the camera being about 9.4 ADU, even the 0.001s exposure is well above this as the bias-corrected mean value here was 157 ADU). The difference in the SNR comes down to the photon count per pixel here and will follow a sqrt(N) function. That function looks like this here in Figure 2-2:  **Figure 2-2**  https://www.cloudynights.com/images/craig-starks/part-5/SimpleSNRvsPhotons.png  The more photons you put in that well to measure, the higher the SNR, even when read noise is zero (blue line). Read noise tacks on an additional penalty (green line) and skyglow does even more (orange line). Now, once the SNR gets above a certain level, our eye perceives it as a relatively clean image and it's not like doubling the SNR will always make an image look "twice as good". But, people do need to realize that getting a decent number of photons into the well is important for a clean image, regardless of how much spatial information is in there. ***Pixel SNR matters for extended objects.***  One other thing we can get from Figure 2-2 here is why read noise can be so important for things like line filter work. The read noise is the only difference between the green and blue lines and both of these assume there is no skyglow (not too far from the truth for narrowband imaging). If you're trying to reach a certain SNR to make the image look reasonable to the eye, it's going to take a lot more photons to do so with the read noise than without. Here, the read noise is a significant part of the noise. The target photon count and skyglow photon count are both low, so they're not contributing very much and the read noise is a constant penalty. Add some skyglow though, and it quickly swamps out the read noise. In the red line, I've added in a decent bit of skyglow but have removed the read noise. As you can see the difference between it the same thing with read noise is approximately nil. An Imaging Example Back in [Part 4](http://www.cloudynights.com/item.php?item_id=2042), I showed some images from Mark Keitel showing an f/7.8 vs. f/5.9 example that, to my eye, showed a real win in SNR for the f/5.9 shot. That was a nice, controlled example. Here, I'll present data with less control, but that some may find striking nonetheless.  A year or so ago, I went out and shot some 5-minute test frames in H-alpha of the Horsehead with an 8" f/5 Antares Newtonian (with a Paracorr) and with my 4" f/4 Borg 101 ED APO. They were shot one right after the other on my QSI 540. Now, the Newt was running at about 1150 mm of focal length (the Paracorr adds 15%), and the refractor was at 400 mm. The Newt gathers 4x as many photons and should have more "information" if aperture is all there is to this. No post-processing was done other than simple stretching. The two images are here in Figure 2-3.  **Figure 2-3**  <https://www.cloudynights.com/images/craig-starks/part-5/HH_Comparison1.jpg>  Now, I don't know about you, but I'm seeing a lot more detail (or at least as much) in the little 4" f/4 scope than I am in the 8" f/5.75. It's also a lot cleaner. It's one I'd share with someone rather than one I'd go back and re-shoot (e.g., with a longer exposure duration to compensate for the change in the photon flux).  OK, so what is going on here? Well, let's stick a few numbers on this (note, these are updated from my blog entry from back then to fix a few things). The 8" Newt does capture 4x as many photons through it's front end, but it's not like they all hit the CCD. The 92% reflectivity mirrors make it such that only 85% of the light gets into the focuser drawtube. Well, that minus a bit for the central obstruction. Toss in some light loss in the Paracorr and you're down to about 78% of the photons hitting the CCD (quick guess of 98% transmission for each of the elements inside the Paracorr). The Borg won't be perfect either. It's a doublet with a reducer/flattener to get it to f/4 and it'll run at about 92% total throughput. Run the math through here and when we account for the focal length differences as well, we get to the answer that each CCD well is getting hit by only 35% as many photons with the Newt ahead of it relative to with the refractor ahead of it. Put another way, the little 4" f/4 is cramming 2.88x as many photons in each CCD well. Looking back at Figure 2-2, it's no wonder the image looks cleaner.  Can this be true? Well, we can measure the same area in both images. My camera's bias signal is about 209 in this area. I measured the mean intensity in a 10x10 region using Nebulosity's Pixel Info tool for three areas right around and in the Horsehead. On the Borg, they measured 425, 302, and 400. On the Newt, they measured 287, 254, and 278. Now, if we pull out the 209 for the bias signal we have 216, 93, and 191 vs. 78, 45, and 69. If we calculate the ratios, we have 2.76x, 2.07x, and 2.77x. Average these and we're at 2.5x. The back of envelope math said it should be 2.88x. That's pretty darn close for envelope math. But, this is again, just simple geometry. We're covering more sky per pixel on the 400 mm scope than on the 1150 mm scope and our aperture, while bigger, hasn't scaled enough to compensate. Guess what would have scaled enough to compensate? An 11.3" f/4 scope.  Again, it's not that the 4" APO will always win. It'll lose out on maximal spatial detail here as 400 mm is undersampling even my skies. A 100" f/4 scope won't produce the same image as a 1" f/4 scope. The target will be a lot bigger and you'll have more detail. But, the brightness (density, photon count, ADU off your CCD, etc.) for an extended object will be the same.***Photon count for an extended object is driven by f-ratio. Image scale is driven by focal length. Want more resolution at the same pixel-wise SNR? Boost the aperture but keep the same f-ratio. Want more SNR in your images and you're either willing to trade some spatial information OR you're already asking for more spatial resolution than your conditions will give? Drop the f-ratio.*** My skies won't support running more than about 1500 mm on even the best of nights. Most nights, I won't see real resolution improvements on this versus 1000 mm. In addition, a sharp but noisy image is unlikely to impress and running at lower f-ratios will pack more photons onto that CCD well boosting the accuracy of our estimate of that patch of sky's true value, my measure of SNR. Summary To re-iterate - as I raise the focal length up to, call it about 1500 mm, I will be increasing the potential spatial information in my images (with my skies, and my 7.4u pixels - your value will vary). Going beyond that and I won't be gaining much, if any, spatial information. Ideally, I'd run each of these focal lengths with as low an f-ratio as possible. Of course, lowering the f-ratio for a constant focal length means increasing the aperture, leading to the notion that aperture rules and leading to the conclusion that we should reach for the biggest scope we can find. But, there are two other sides to this that should be considered. First, as we go up in focal length, passing that point into over-sampling is leading you to lose SNR on extended objects without gaining any actual spatial information. Second, spatial information alone may not be what we're all after. I will gladly trade some amount of spatial information for a cleaner galaxy or more galaxies in my background. For that, photon density per CCD well rather than total number of photons collected matters most. Here is where the f-ratio clearly steps in as it is what determines the photon density (aperture = total photons; focal length = spread of photons; f-ratio = density of photons).  Many amateurs routinely run in photon-poor conditions. We use small pixels on DSLRs. We use one-shot color cameras with their built-in filters or we add our own filters that cut the photon counts (especially, but not limited to line-filters). We grab f/8 and f/10 scopes that spread the photons thin and we image for a minute or so. All of these conspire to cut the photons per pixel down which cuts the SNR of our extended objects. Run that scope at a lower f-ratio (which, yes, will make each pixel cover more sky as that's the whole point) and you'll image that nebula better. Myth 3: There is something about film that made f-ratio apply that isn't true of CCDs (or there is something about CCDs that makes the f-ratio not apply). Hopefully, the above is at least making it clear that the f-ratio does matter with CCDs for our DSOs. In some ways, this needn't be covered here, but it is something I think it's worth setting the record straight on. Much has been made about "reciprocity" in film. First, let's get the terms clear.  "[Reciprocity](http://en.wikipedia.org/wiki/Reciprocity_(photography))" refers to the reciprocal relationship between exposure duration and intensity.That is, that the density on film is equal to the exposure time multiplied by the intensity.Inherent in this is the notion that film is a linear medium while inside a certain range."Reciprocity failure" is the breakdown of this.With very low flux rates, film became non-linear.It takes several dozen photons to get a halide crystal to form a latent image and if they don't arrive within a certain amount of time, the captured photons are lost.This, of course, hits astrophotographers (who use film) extensively.But, this has nothing to do at all with the notion from general photographers that doubling the exposure duration is equal to one f-stop.These photographers quite correctly used the reciprocity rule as they had enough photon flux to ignore reciprocity failure.  Now, the film response isn't purely linear and it's not always easily characterized, but it certainly can be quite linear over a certain range.Here, in Figure 3-1, we have a plot of Kodak Plus-X film at various exposure durations and at three different development times.  **Figure 3-1**  https://www.cloudynights.com/images/craig-starks/part-5/film_response.png  The "knee" on the left there is the reciprocity failure, but beyond -2.0 or so on the graph, we're looking nice and linear.Film can be linear and the use of a reciprocal relationship between exposure duration and f-ratio by photographers isn't the result of any oddness of film.Rather, it's the result of the fact that for much of this graph, things are linear and we can swap out one for the other.That is the definition of reciprocity.Again, where it breaks down is in the very low flux situations (the left side of this graph), typically only encountered by astrophotographers, high-speed photographers, and microphotographers (all with low flux counts). In this reciprocity failure area, you fail to record the target. So, getting the flux rate above this is crucial to recording the target. Lowering the f-ratio will, of course, get you up off this knee better. But, f-ratio here is helping you get out of the non-linear zone. The golden rule of f-ratios changing the density of the recorded image has nothing to do with this zone and has everything to do with the nice linear zone. One f-stop is equivalent to doubling the flux, a relationship that only holds when things are linear.  Now, the big difference between CCD and film here (apart from an overall difference in sensitivity) is the fact that the CCD response is not only linear in the meat of the range, but it is also linear on the low end. Here is a plot from Kodak's spec sheet for one of their sensors. Numerous similar examples exist on the web (e.g., a [Kodak KAF-1602E's response from an SBIG ST8E](http://www.jca.umbc.edu/telescope/Glossary/CCD_SBIG_ST-8E.html)measured by the Astrophysics and Space Science Institute at UMBC, and a [sample spec sheet from Apogee](http://www.ccd.com/ccd105.html)):  **Figure 3-2**  https://www.cloudynights.com/images/craig-starks/part-5/linearity.png  You'll notice that the bottom end is nice and linear. The top end starts to show saturation (much as film does). But again, it is a misnomer to say that "CCD's are linear and film isn't." Film can be quite linear and typically is very linear within a certain range (the range photographers typically use and the range the f-ratio rule is made for). Both show "reciprocity" and that's a good thing. Above and below that range it goes non-linear. For many of our CCDs, above a certain range, they also go non-linear. The great thing is that not only do they not suffer from reciprocity failure, but they are far more sensitive overall than film. However, in the linear range, doubling the photon flux (e.g., by dropping by one f-stop) doubles the density on the film and doubles the number of photons captured by the CCD.  With this, we have likely reached the conclusion of this series on SNR. Others may wish to continue with topics like the role that f-ratio plays on star shape, the role for imperfections in CCDs, a more thorough treatment of the role of f-ratio on stellar sources, or a more thorough treatment of object SNR (or, perhaps more appropriately, a treatment of "image quality" - how well all of the information from that portion of the sky is being recorded.) I hope from my modest efforts on the topic, that some have learned a few things along the way. Whether you enjoyed the ride, found it obtuse and confusing or thought it myopic, I wish you clear, steady, and dark skies so that you can enjoy fishing for those photons. |

**Source:** [**http://www.stark-labs.com/craig/articles/articles.html**](http://www.stark-labs.com/craig/articles/articles.html)

**Articles posted on Cloudy Night Forum.**